Mag. Otto Wurnig
Institut für Mathematik
Universität Graz
Heinrichstraße 36

Von der Binomial- zur Normalverteilung mit Hilfe des Computers

1. Wahrscheinlichkeitsverteilung im Lehrplan der AHS

Das Kapitel "Wahrscheinlichkeitsrechnung" wurde in Österreich das erste Mal im Schuljahr 1970/71 in den AHS nach dem reformierten Lehrplan in der 8. Klasse unterrichtet[1]. Es standen dafür keine approbierten Lehrbücher zur Verfügung. Da das Kapitel "Kombinatorik" bereits im 19.Jahrhundert in den Gymnasien einen Teil des Mathematiklehrplans ausmachte, ist es nicht weiter verwunderlich, daß vor allem klassische Wahrscheinlichkeitsrechnung auf kombinatorischer Grundlage gebracht wurde. Die "Binomische Verteilung" und die "Normalverteilung" war dabei nur für das Realgymnasium vorgesehen, wobei sie dort anfangs wegen Zeitmangels kaum durchgenommen werden konnten. Die letzte Reifeprüfung nach diesem Lehrplan fand 1981 statt.

Im Mathematiklehrplan 1978[2] (1. Reifeprüfung 1982, letzte 1992) wurde die Wahrscheinlichkeitsrechnung auf die 7. und 8. Klasse aufgeteilt. Für die 7. Klasse wurden die Grundlagen bis zur Formel von Bayes, für die 8. Klasse die Wahrscheinlichkeitsverteilungen vorgesehen. Approbierte Lehrbücher standen dieses Mal von Anfang an zur Verfügung und Umfragen in den didaktischen Lehrveranstaltungen am Insitut für Mathematik der Universität Graz ergaben noch Anfang der 90er-Jahre, daß sich fast alle von der AHS kommenden Studenten an Beispiele zur Binomialverteilung erinnerten, aber nur ca. 50% auch an Aufgabenstellungen zur Normalverteilung !

Im neuen Lehrplan 1989 [3](1. Reifeprüfung 1993) verteilt sich die Wahrscheinlichkeitsrechnung wiederum auf die 7. und 8. Klasse. Aber bereits die ersten drei Sätze zur Wahrscheinlichkeitsrechnung im Lehrplan zeigen, daß sich Inhalt und Methode deutlich verändert haben !
"Schwerpunkt soll das Arbeiten mit zumindest einer Wahrscheinlichkeitsverteilung und das Bearbeiten von Problemen der Beurteilenden Statistik sein. Dabei ist eine ausführliche Behandlung des Berechnens von (bedingten) Wahrscheinlichkeiten einzelner Ereignisse nicht unbedingt erforderlich. Die Verwendung von Rechengeräten und geeigneter Software ist zweckmäßig."
Der Lehrplan läßt fürs erste offen, welche Verteilungen in der 7. Klasse durchgenommen werden sollen. Es wird nur etwas später betont, daß "Anwendungsaufgaben mit der Binomial- oder Normalverteilung" zu lösen sind. In der 8. Klasse ist dann das "Bearbeiten von Problemen (etwa Berechnen von Wahrscheinlichkeiten, Schätzen, Testen) mit bekannten oder auch neuen Verteilungen" vorgesehen. Allenfalls kann ein "vertieftes Betrachten von Wahrscheinlichkeitsverteilungen: Etwa: Vergleichen von Wahrscheinlichkeitsverteilungen (beispielsweise hinsichtlich ihrer Anwendbarkeit)" angestrebt werden.

2. Wahrscheinlichkeitsverteilung im Spiegel der Lehrbücher.

In Österreich wurden in der Gratis-Schulbuchaktion für das Schuljahr 1992/93 vier approbierte Lehrbücher für die 5.-8. Klasse zur Auswahl angeboten. Davon stammt eines vom Reniets Verlag und zwar das Buch "Novak, Mathematik Oberstufe 1-4" und drei vom Verlag Hölder-Pichler-Tempski. Es sind folgende Lehrbücher:
Bürger-Fischer-Malle, Mathematik Oberstufe 1-4 (kurz: B-F-M),
Reichel-Müller-Hanisch, Lehrbuch der Mathematik 5-8 (kurz: R-M-H),
Szirucsek-Dinauer-Unfried-Schatzl, Mathematik 5-8 (kurz: S-D-U-S).

Da nur das Lehrbuch von Novak in der 7. Klasse Aufgaben zur Binomial- und Normalverteilung enthält, soll dieses Buch vorangestellt werden.

Nach zwanzig Seiten Einführung in die Wahrscheinlichkeitsrechnung und Statistik wird in der Lösung zu einem Beispiel der folgende Satz für Binomialverteilungen ohne Beweis angeboten:
"Wenn ein Erfolg E in einem n-mal unter denselben Bedingungen durchgeführten Versuch der Wahrscheinlichkeit p eintritt, dann ist die Wahrscheinlichkeit P(X=k) für genau k Erfolge in n Versuchen

$$P(X = k) = \begin{bmatrix} n \\ k \end{bmatrix} p^{k} (1 - p)^{n - k} \quad ."[4]$$

Nach zwei Seiten Aufgaben und einer Seite einführender Betrachtungen wird auf einer Seite die Normalverteilung aus der Binomialverteilung wie folgt hergeleitet[5]:
"Wir stellen uns die Wahrscheinlichkeitsverteilung der Binomialverteilung nun als stetige Funktion von $x \epsilon R$ vor:

$$f(x) = P(X = x) = \begin{bmatrix} n \\ x \end{bmatrix} p^{x} (1 - p)^{n - x} \quad .$$

Nehmen wir an, die Anzahl der Nagelreihen n (in der Praxis gleichbedeutend mit der Anzahl der Betriebseinflüsse) wird sehr groß; dann erhalten wir für μ=n.p und σ=$\surd$(n.p.(1-p)) durch geeigneten Grenzübergang die Funktion:

$$f(x) = \frac{1}{\sigma \surd(2 \pi)} \hat{e}^{- (x - \mu)^2 / (2 \sigma^2)}$$

Diese Funktion wird Dichtefunktion der Normalverteilung genannt. Die Wahrscheinlichkeit, daß die Zufallsvariable X einen in einem bestimmten Intervall liegenden Wert annimmt, entspricht der in diesem Intervall zwischen der x-Achse und f(x) eingeschlossenen Fläche[2]." In der Fußnote 2 wird begründet, daß auf einen Beweis verzichtet wurde, da er Kenntnisse voraussetzt, die über die Lehrplanforderungen hinausgehen.
In dieser Darstellung bleiben beide Formeln für den Schüler Zauberformeln, mit denen er zwar rechnen kann, deren Zustandekommen aber für ihn rätselhaft bleiben muß. Leider wird auch im Buch für die 8. Klasse keine weitere Erklärung gegeben. Am Ende des Kapitels wird in der 7. Klasse auf die Unterstützung durch den Computer in einigen Zeilen hingewiesen, wobei besonders betont wird, daß viele Beispiele des Kapitels Wahrscheinlichkeitsverteilungen mit SuperCalc5 bearbeitet werden könnten.

Alle drei anderen Lehrbücher stellen aus durchaus verständlichen Gründen in der 7. Klasse die Binomialverteilung in den Mittelpunkt. In der 8. Klasse wird dann die Normalverteilung eingeführt. Im folgenden sollen die unterschiedlichen Konzepte für die Einführung der Normalverteilung in der 8. Klasse kurz skizziert werden:

Im Lehrbuch R-M-H 8 wird nach der Einführung der stetigen Zufallsvariablen der Begriff und die Eigenschaften der Wahrscheinlichkeitsdichtefunktion gebracht und anschließend Erwartungswert und Streuung mit Hilfe der Integralrechnung definiert. Das nächste Kapitel stellt die Definition der Normalverteilung, die als die wichtigste stetige Verteilung bezeichnet wird, an die Spitze.

"Die durch die Wahrscheinlichkeitsdichtefunktion

$$f: \quad y = \frac{1}{\sigma \sqrt{(2\pi)}} \, e^{-0.5\,((x-\mu)/\sigma)^2}$$

festgelegte Verteilung heißt Normalverteilung mit den Parametern μ und σ, kurz $N(\mu;\sigma^2)$-Verteilung. Der Graph von f heißt Gaußsche Glockenkurve.
Ist insbesondere $\mu=0$ und $\sigma=1$, so bezeichnen wir die Wahrscheinlichkeitsdichtefunktion statt mit f mit phi und sprechen von der Standardnormalverteilung N(0;1)."[6]

Es folgen zehn Seiten, auf denen die neuen Begriffe erklärt, und das Arbeiten mit der Verteilungsfunktion der Standardnormalverteilung

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{(2\pi)}} \, e^{-t^2/2} \, . \, dt$$

mit Hilfe der Tabelle für das Lösen verschiedenster Anwendungsaufgaben geübt wird. Dann erst ist die Approximation der Binomialverteilung durch die Normalverteilung vorgesehen. Die Wichtigkeit dieser Näherungsformel wird durch ein eigenes Kapitel "Systematisches Lösen von Anwendungsaufgaben zur Binomialverteilung mittels der Normalverteilung"[7] unterstrichen.

Die vier Grundaufgaben werden sehr gut in Musteraufgaben herausgestellt. Leider wird dabei auf eine anschauliche Interpretation mit Hilfe der Glockenkurve vergessen. War der Einstieg zur Normalverteilung doch relativ abstrakt, so werden an vielen Beispielen die Schüler in das richtige Verwenden der Formeln und der Tabelle eingeschult. Allerdings besteht die Gefahr, daß nach dieser Art der Einführung beim Rechnen der Zusammenhang zur Theorie beiseite geschoben werden könnte.

In Lehrbuch S-D-U-S 8 werden zuerst diskrete Verteilungen behandelt, wobei außer der von der 7. Klasse her bekannten Binomialverteilung, auch die Hypergeometrische Verteilung und die Poisson-Verteilung drankommen. Nach einer Einführung in die Grundbegriffe von stetigen Verteilungen auf sechs Seiten folgt ein anschaulicher Einstieg in die Normalverteilung an Hand eines Beispiels über das Körpergewicht Neugeborener. Dabei ergibt die konkrete Auswertung der Tabelle der Geburtsgewichte von tausend im LKH Graz im Jahre 1991 lebend geborenen Mädchen als Graph der Dichtefunktion eine Glockenkurve. Die e-Funktion dafür wird dann einfach hingeschrieben. Allerdings sollen die Schüler in den nachfolgenden Beispielen[8] einige Überlegungen zur Funktion anstellen, wobei sie das auf der Lehrbuchdiskette befindliche Programm DGAUSS.EXE zum Zeichnen von Glockenkurven zu verschiedenen Werten für μ und σ verwenden sollen. Sechs Seiten später wird dann die Binomialverteilung durch die Normalverteilung angenähert.

Im Lehrbuch B-F-M 4 wird der direkte Weg über die Binomialverteilung eingeschlagen. Nach einer kurzen Wiederholung der Berechnung des Erwartungswertes μ und der Varianz σ einer binomialverteilten Zufallsvariablen H wird auf die Problematik der Berechnung von Wahrscheinlichkeiten der Form $P(H \leq k)$ bzw. $P(H \geq k)$ mit Hilfe einer Tabelle hingewiesen. Da derartige Tabellen prinzipiell nur für eine begrenzte Anzahl von Werten von n hergestellt wird, benötigt man eine Methode, mit der solche Wahrscheinlichkeiten für große n näherungsweise bestimmt werden kann.

Ausgangspunkt der Überlegungen sind vier im Buch als Histogramme bildlich dargestellten spezielle Verteilungen, zu denen im Buch auch die Tabellen mitgeliefert werden. Aus der Darstellung sollen die Schüler erkennen, daß Wahrscheinlichkeiten in Stabdiagrammen durch Stablängen, in Histogrammen durch Flächeninhalte (von Rechtecken mit der Klassenbreite 1) dargestellt werden können.
Um diese Histogramme besser vergleichen zu können, wird im Buch in einer weiteren Figur - siehe Fig. 5.a-d - jedes Staffelbild über eine gemeinsame Skala, die "z-Skala" gelegt. Dabei müssen die ursprünglichen Rechtecke verändert werden. Die Breite 1 wird durch σ dividiert und die Länge mit σ multipliziert, wodurch der Flächeninhalt eines jeden Rechteckes unverändert bleibt.

$n=50,\ p=0{,}5,\ \mu=25,\ \sigma=3{,}54$

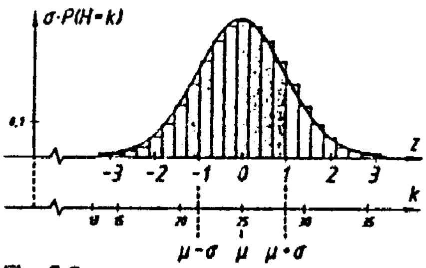


Fig. 5.2 a

$n=50,\ p=0{,}3,\ \mu=15,\ \sigma=3{,}24$

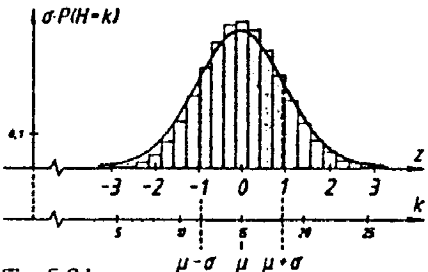


Fig. 5.2 b

$n=100,\ p=0{,}5,\ \mu=50,\ \sigma=5$

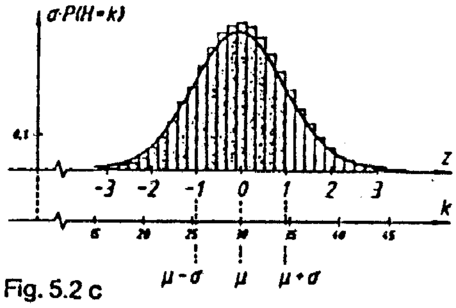


Fig. 5.2 c

$n=100,\ p=0{,}3,\ \mu=30,\ \sigma=4{,}58$

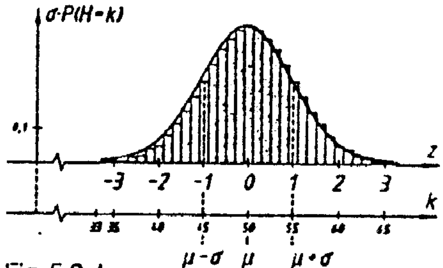


Fig. 5.2 d

Mit Hilfe von Berechnungen an desen Histogrammen soll der Schüler zum Schluß kommen, daß die Wahrscheinlichkeit $P(H \leq \mu+z.\sigma)$ für alle binomialverteilten Zufallsvariablen H annähernd denselben Wert hat und somit von n,p, aber auch von μ,σ unabhängig ist. Sie hängt nur von z ab. Weiters soll der Schüler vermuten, "daß die Wahrscheinlichkeit $P(H \leq \mu+z.\sigma)$ annähernd gleich dem Inhalt der Fläche ist, den die (im Buch rot eingezeichnete) Kurve mit der ersten Achse im Intervall $]-\infty,z]$ einschließt"[9].

Bei dieser Kurve handelt es sich um den Graphen der Funktion phi mit:

$$\mathrm{phi}(x) = \frac{1}{\sqrt{(2\pi)}}\, \hat{e}^{-t^2/2}, \qquad x \in R$$

Man bezeichnet sie als Gaußsche Glockenkurve und den Inhalt der Fläche, den diese Kurve mit der 1. Achse im Intervall $]-\infty, z]$ einschließt mit $\Phi(z)$.
Nach dieser ausführlichen Darlegung erfolgt drei Seiten weiter die Definition einer normalverteilten Zufallsvariablen: "Eine Zufallsvariable X heißt normalverteilt mit den Parametern μ und σ (wobei $\mu, \sigma \in R$ und $\sigma > 0$), wenn für alle $z \in R$ gilt: $P(X \leq \mu + z.\sigma) = \Phi(z)$"[10].

In diesem Konzept steht somit der Übergang von einer diskreten Verteilung zu einer stetigen im Mittelpunkt. Man versucht dies durch Staffelbilder veranschaulichen. Da aber für n=50 bzw. 100 die Staffelbilder schwer zu zeichnen sind, mußten verschiedene Einheiten auf den Achsen gewählt werden. Dadurch geht ein wichtiger Teil der Bildwirkung verloren, und der Schüler kann den ausführlich beschriebene Übergang vom Stab zum Rechteck in der angebotenen Zeichnung nur sehr schwer nachvollziehen.
Beim Standardisierungsprozeß der Normalverteilung, wird die Glockenkurve bewußt unverändert gelassen, und nur die Skalen auf beiden Achsen geändert. Mitautor Malle meint im Kommentar Oberstufenlehrplan[11], daß das "Kurvenschieben" graphisch sehr umständlich sei. Es sei für die Schüler leichter verständlich, wenn stets beide Skalen - die x-Skala und die z-Skala - gezeichnet werde. Weiters wurde im Buch bewußt, die in den anderen Büchern verwendete Transformationsgleichung $z=(x-\mu)/\sigma$ vermieden. Ob dieses Konzept wirklich geschickt ist, wird erst im Unterricht zu überprüfen sein.
Die e-Funktion für die standardisierte Glockenkurve "fällt" zwar hier auch "vom Himmel". Ihre Zweckmäßigigkeit kann aber vom Schüler jederzeit an der (roten) Begrenzungskurve überprüft werden.

Durch das Nebeneinanderstellen der vier Lehrbuchkonzepte sollten die verschiedenen Vorgangsweisen mit ihren Stärken und Schwächen aufgezeigt werden. Aus verständlichen Gründen ist in allen besprochenen Büchern der Computer, wenn überhaupt, nur am Rande erwähnt bzw. zum Einsatz vorgeschlagen. Über weite Strecken sind die Vorschläge so gestaltet, daß der Mathematiklehrer frontal, so gut und rasch wie möglich, die wichtigsten Fakten und Formeln vermittelt, und der Schüler dann jenen Teil, den er wirklich zum Lösen der Aufgaben benötigt, auswendiglernt und verwendet.
Die Formel für die Normalverteilungung "fällt" in allen Büchern "vom Himmel". Kenntnisse der Integralrechnung werden eigentlich nicht benötigt, da alle Aufgaben mit Hilfe der $\Phi(z)$-Tabelle gelöst werden können und auch Überlegungen für $\Phi(-z)$ später oft über Formeln - ohne Kontakt mit der Glockrnkurve - vorgenommen werden. Es scheint den Lehrbuchautoren in erster Linie wichtig zu sein, daß schnell Aufgaben gelöst werden können. Tatsächlich ist die Zeit in der 8. Klasse nur kurz bemessen.

Leider kommt aber bei einem solchen Unterricht die Entwicklung allgemeiner mathematischer Fähigkeiten zu kurz. Nach dem neuen Lehrplan 1989 sind im Zusammenhang mit dem Erwerb von mathematischen Wissen und Können und dem Anwenden von Mathematik folgende Lernziele anzustreben:

* Argumentieren und exaktes Arbeiten,
* Darstellen und Interpretieren,
* Produktives geistiges Arbeiten,
* Kritisches Denken.

In meinen computerunterstützten Unterrichtskonzept, das in sechs Kapitel erläutert wird, werden diese Ziele mitberücksichtigt. Das Konzept habe ich 1992/93 in der 8.d Klasse am BRG Graz Keplerstraße durchgeführt und auch die schriftliche und mündliche Reifeprüfung computerunterstützt gestaltet.

3. Neuerungen während des ersten Durchgangs der Oberstufenreform 1989-1993

Mit Beginn des Schuljahres 1989/90 begann in Österreich die reformierte Oberstufe mit neuen Rahmenbedingungen:
Die Klassenschülerhöchstzahl wurde von 36 auf 30 herabgesetzt, was 1989/90 kleinere 5. Klassen bewirkte. Am Beginn der Oberstufe gebildete Klassen müssen erst dann am Ende eines Schuljahres aufgelöst werden, wenn die durchschnittliche Schüleranzahl pro Klasse eines Jahrgangs 14 unterschreitet. Dies hatte zur Folge, daß im Schuljahr 1992/93 ein Großteil der 8. Klassen an AHS in Österreich die Schülerzahl 20 unterschritt.
Außerdem verfügen seit 1990/91 alle AHS-Langformen über einen Raum mit vierzehn Computern, so daß später im Schuljahr 1992/93 in der 8. Klasse mehr als die Hälfte der Schüler ein eigenes Gerät zur Verfügung hatten. Damit sind die Voraussetzungen für einen computerunterstützten Unterricht gegeben.
Im Rahmen der Oberstufenreform wurden erstmals Wahlpflichtgegenstände angeboten. Bei der Wahl entscheiden sich seit Jänner 1990 Jahr für Jahr so viele Schüler in Österreich für das Fach Informatik in der 6.-8. Klasse, daß dieser Gegenstand österreichweit an der Spitze aller Wahlpflichtgegenstände liegt. Fast jede Oberstufenklasse besitzt daher einige Schüler, die die Lehrer anderer Gegenstände, falls sie in den Computerraum gehen wollen, bestens unterstützen können.
Es gilt seit 1989/90 ab der 5. Klasse aufsteigend ein neuer Mathematiklehrplan[3]. Dieser sieht in der Bildungs- und Lehraufgabe und in den didaktischen Grundsätzen den Personalcomputer als mögliches Rechengerät nicht nur ausdrücklich vor, sondern fordert ein damit "vertraut werden".
Aus den angeführten Gründen besteht von den Mathematiklehrern an AHS ein Bedürfnis nach passender Software. Im Jahr 1990 erhielten sie österreichweit die Tabellenkalkulation SuperCalc5 und das Stochastikprogramm von W. Stormer, im Juni 1991 das Computeralgebrasystem DERIVE, alle mit Schullizenz. In vielen Fortbildungsveranstaltungen zu diesen Produkten wurde außerdem als willkommene Ergänzung das Shareware-Programm MATHEASS vorgestellt.

Mit den genannten Programmen wurde eine brauchbare Basis für den Mathematikunterricht geschaffen. Diese muß ständig durch Updatings auf die neuen Versionen und durch die Weitergabe selbsterstellter Lehrer-Programme erweitert werden.

4. Die Bedingungen für die 8.d Klasse am BRG-Kepler.

In die 5.d Klasse gingen 1989/89 25 Schüler, in die 7.d Klasse 1991/92 19 Schüler und in die 8.d Klasse 1992/93 noch 16 Schüler. Bis auf drei Schüler hatten alle von der 6. bis 8. Klasse den Wahlpflichtgegenstand Informatik gewählt.
Im Herbst 1991 begann ich immer häufiger Mathematikstunden computerunterstützt im 14-Geräteraum zu halten. Ich verwendete dabei alle vorher genannten Computerprogramme, mußte aber bald erkennen, daß DERIVE einigen Schülern im Handling große Probleme bereitete, aber gerade dieses Programm große Möglichkeiten für eine Umgestaltung des Unterrichtes bietet.
Daher versuchte ich durch eine intensivere Einführung in DERIVE diesen Schülern zu helfen, mußte aber bald erkennen, daß die nötige Fertigkeit zu selbständigen Arbeiten mit DERIVE in der Schule allein nicht zu erreichen ist. Da aber anderseits einige Schüler sich bald recht gut mit DERIVE auskannten, die Hausübung in Mathematik zu Hause am eigenen Computer zu machen begannen, beschloß ich im Jänner 1992, ab sofort die Schularbeiten alternativ mit und ohne DERIVE anzubieten. Die Schüler hatten dabei eine Woche vor dem Schularbeitstermin ihre Wahl zu treffen. Wollten bei der 2. Schularbeit im Jänner 1992 nur 9 von 19 die Schularbeit cmputerunterstützt schreiben, so waren es bei der 5. Schularbeit im Juni 1992 bereits alle Schüler.
Am Beginn der 8. Klasse brachten alle Schüler die schriftliche Einverständniserklärung der Eltern mit, daß ab sofort bei allen Hausübungen und Schularbeiten DERIVE verwendet werden könne. Bis Ende Oktober hatten alle einen PC daheim zur Verfügung, drei davon ein älteres Leihgerät der Schule.

Da die Schülerzahl nur mehr sechzehn betrug, konnten in der Mathematikstunde praktisch alle Schüler im Informatikraum an einen eigenen Computer arbeiten. Aus organisatorischen Gründen ergab sich eine 1:1 Aufteilung zwischen den Zeiten im Klassen- und Computerraum.

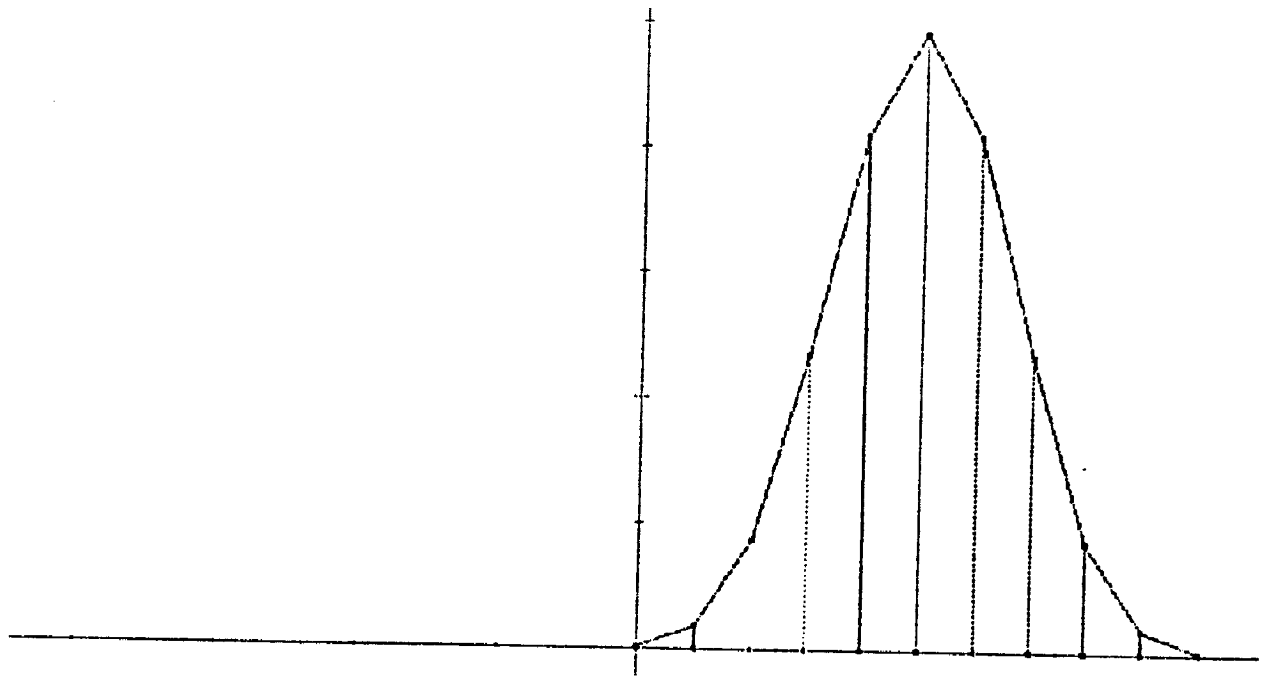
5. Mein Konzept zum Thema : Von der Binomial- zur Normalverteilung

Da in der 7. Klasse mit der Binomialverteilung bereits Beispiele gelöst worden sind, habe ich die Normalverteilung als Grenzverteilung der Binomialverteilung für große n und k eingeführt. Bei dieser Gelegenheit wurde die Problematik des Überganges von einer diskreten zu einer kontinuierlichen Verteilung ausführlich diskutiert. Es soll aber auch darauf hingewiesen werden, daß stetige Verteilungen nicht nur als Grenzwert von diskreten Verteilungen, sondern auch unabhängig davon von Bedeutung sind.
Die bei Grenwertbetrachtungen zur Binomialverteilung entstehende Glockenkurve kann jedoch nur dann als Graph einer Funktion des Typs e^{-x^2} erkannt werden, wenn dies im vorangestellten Kapitel über Exponentialfunktionen entsprechend vorbereitet wurde. Um bei stetigen Verteilungen den Flächeninhalt als Maß für die Wahrscheinlichkeit erkennen zu können, hat man die Integralrechnung mit Flächeninhaltsberechnungen vorher durchzunehmen.
Üblicherweise werden für die Berechnung von Integralen von e^{-x^2} Tabellen verwendet, Setzt man stattdessen den PC und eine geeignete Software ein, so wird klarer, daß es sich dabei um die numerische Berechnung bestimmter Integrale handelt. Damit ist eine Methode entwickelt, die auch auf andere stetige Verteilungen übertragbar ist. Mit Hilfe solcher Software können Beispiele zur Normalverteilung statt mit einer Methode (Tabelle) mit mehreren gelöst werden.

Ich stelle in den folgenden fünf Punkten die wesentlichen Schritte des von mir durchgeführten computerunterstützten Konzeptes vor.

5.1 Stabdiagramm-> Histogramm-> Wahrscheinlichkeitsdichtepolygon

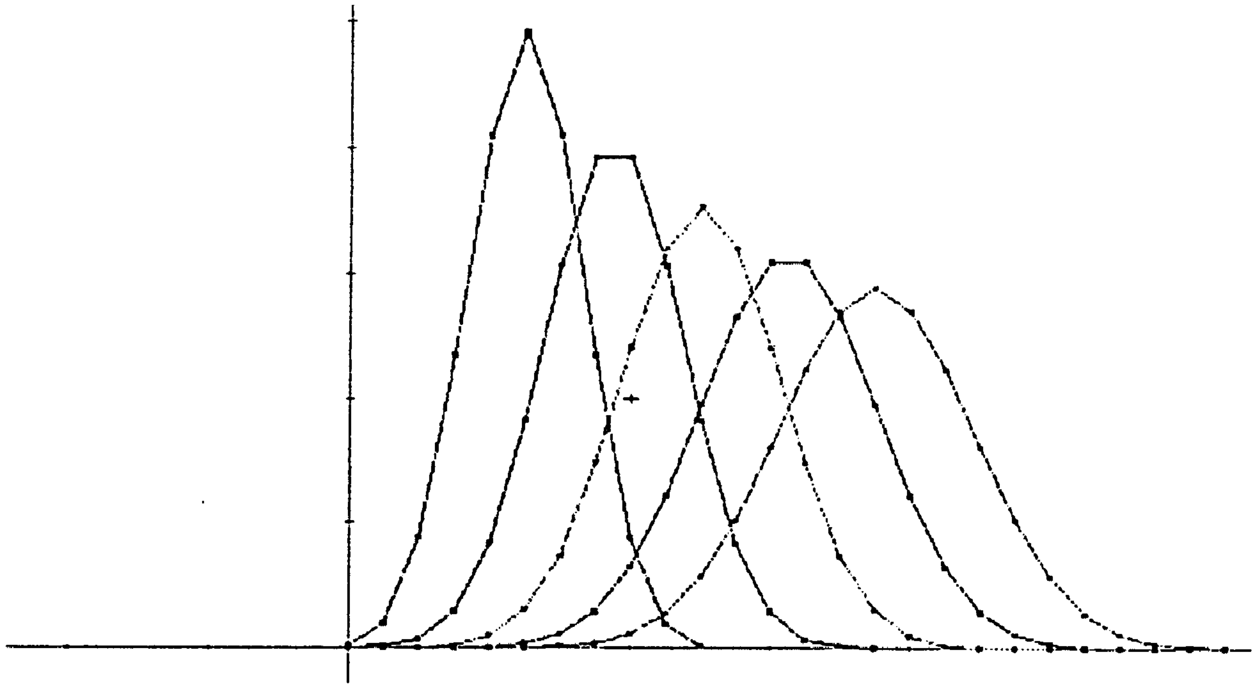
Im Klassenraum stellte jeder Schüler für sich die Tabelle der Binomialverteilung für n=10 und p=0.5 auf und verglich sie mit dem Ergebnis am Demonstrationscomputer. Dann erstellte er dazu in einer Zeichnung das Punkt- und Stabdiagramm, mit der Breite 1 das Histogramm und schließlich das Wahrscheinlichkeitsdichte-Polygon(W-Polygon).
Im Computerraum wird das nächste Mal die Zeichnung (Punkt-, Stabdiagramm und W-Polygon) für p=0.5 und p=10 mit DERIVE durchgeführt.



Lernziele:
Der Schüler soll erkennen,
* daß Wahrscheinlichkeiten in Stabdiagrammen durch Stablängen, in Histogrammen bzw. W-Polygonen durch Flächeninhalten dargestellt werden,
* daß sich dabei beim Histogramm gegenüber dem Stabdiagramm das Intervall auf der 1. Achse um eine halbe Einheit auf beiden Seiten vergrößert (Stetigkeitskorrektur),
* daß in Stabdiagrammen die Summe der Stablängen 1,
* daß bei W-Polygonen der eingeschlossene Flächeninhalt 1 ist.

5.2 Vom Wahrscheinlichkeitsdichtepolygon zur Verteilungsfunktion

Anschließend wurde für n = 15, 20, 25 und 30 mit der gleichen Wahr scheinlichkeit in eiem Bild die W-Polygone erstellt.



Lernziele:

Der Schüler soll am selbst erstellten Bild erkennen,

* daß der Streckenzug immer mehr zur Kurve wird,
* daß das Bild symmetrisch ist,
* daß μ=n.p immer größer und b(μ) immer kleiner wird,
* daß die Fläche (Inhalt= 1) immer flacher und breiter wird,
* daß im Grenzfall für n->∞ die Fläche zur 1. Achse entartet.

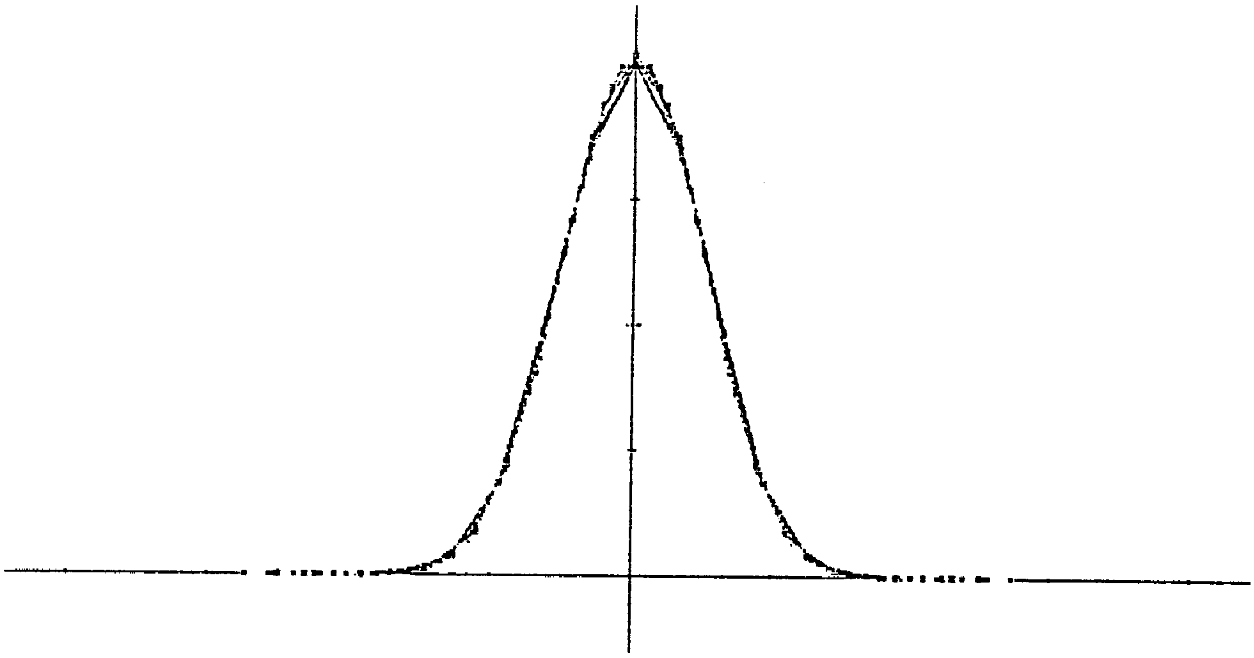
5.3 Die Standardisierung

Da die Kurven nach rechts vom Bildschirm zu verschwinden begannen, schlugen die Schüler selbst eine Zentrierung vor. Die ersten Streckenzüge wurden speziell transformiert, auf die Möglichkeit dafür eine Formel zu findenen (z=x-μ) mußte ich erst hinweisen. Für alle transformierten Polygone war nun μ=0. Daher überlegten wir uns einen gemeinsamen Wert σ. σ=1 wurde von den Schülern selbst vorgeschlagen. Nach einer Diskussion verstanden sie, daß die vom Polygon eingeschlossene Fläche nur dann den Inhalt 1 behalten kann, wenn man gleichzeitig staucht und streckt.

Mit etwas Hilfe erhielten die Schüler die allgemeinen Transformationsgleichungen für die gesamte Standardisierung:

$$z = (x - \mu)/\sigma \quad \text{und} \quad b(z) = \sigma . b(x)$$

Sie führten diese Transformation am Bildschirm mit den vorhandenen W-Polygonen durch, und stellten dabei überrascht fest, daß vor ihren Augen eine Kurve entstand. Diese war um umso klarer zu erkennen, je größer n wurde. Da jedes neue standardisierte Polygon auf einem Farbmonitor mit einer anderen Farbe gezeichnet wurde, ließ sich dieser Vorgang gut verfolgen.



Lernziele:

* Der Schüler soll von jedem W-Polygon die Standardisierung durchführen können.
* Der Schüler soll aus dem Bild erkennen, daß Polygone mit größerem n immer besser auf die entstehende Kurve passen.
* Der Schüler soll aufgrund des Prozesses zur Formulierung folgender Untersuchungsaufgaben angeregt werden:

- Gibt es eine Funktions f, deren Graph die erhaltene Kurve ist?
- Wie groß ist das Maximum dieser Kurve, also f(0) ?

5.4 Die Berechnung des Maximums der Grenzkurve und der anschließend Suchvorgang nach f(x)

Die Formel für die Binomialverteilung lautet:

$$B(k) := COMB(n, k)\ p^{k}\ (1 - p)^{n - k}$$

COMB(n,k) ist die Schreibweise von DERIVE für "n über k".
Die Formeln für den Erwartungswert μ und die Varianz σ sind:

$$\mu := n\ p \qquad \sigma := \sqrt{(n\ p\ (1 - p))}$$

Jetzt wird für n, k und p substituiert: n->2n, k->n, p=1/2 und anschließend vereinfacht.

$$B(n) := COMB(2\ n, n) \left[\frac{1}{2}\right]^{n} \left[1 - \frac{1}{2}\right]^{2\ n - n}$$

$$B(n) := \frac{\left[n - \frac{1}{2}\right]!}{\sqrt{\pi}\ n!}$$

In die Formeln für μ und σ eingesetzt erhält man: μ=n, $\sigma=\sqrt{(2n)}/2$.
b(μ) soll nun mit σ multiplziert und dann vereinfacht werden:

$$\frac{\sqrt{2}\ \sqrt{n} \left[n - \frac{1}{2}\right]!}{2\ \sqrt{\pi}\ n!}$$

Bildet man von diesem Ausdruck den Grenzwert für n->∞, so erhält man das Maximum mit

$$1/\sqrt{(2\pi)} \approx 0.3989.$$

Die Berechnung dieses Grenzwertes bereitet den Schülern wenig Probleme. Aber das Finden der Verteilungsfunktion war nicht einfach.
Zuerst wurde im Klassengespräch erarbeitet, daß eine Funktion mit folgenden Eigenschaften zu bestimmen ist:

* $f(0) = 1/\sqrt{(2\pi)}$ mit $f(x) \geq 0$ für alle x.
* Der Graph von f ist glockenförmig u. symmetrisch zur y-Achse.
* Die von $-\infty$ bis $+\infty$ von der Kurve und der x-Achse eingeschlossene Fläche hat den Inhalt 1.

Die erste Eigenschaft kann keine Potenzfunktion, sondern nur eine Exponentialfunktion besitzen. Die zweite Eigenschaft hat zum Beispiel die Funktion e^{-x^2}, was mit DERIVE einfach mit "Plot Plot" am Bildschirm überprüft werden kann. Die Funktion $f(x)=1/\sqrt{(2\pi)}.e^{-x^2}$ hat dann beide Eigenschaften.

Die Schüler integrierten diese Funktion von $-\infty$ bis $+\infty$ und erhielten mit DERIVE $\sqrt{2}/2 < 1$. Daher wurde folgende Gleichung angesetzt:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)}}\, \hat{e}^{-a x^2}\, dx = 1$$

Der erste Lösungsversuch mißlang. Der Computer gab bei den verschiedensten Versuchen sehr häufig "no solutions found" oder "? = 1" aus !

Da aber die Schüler selber zu einem brauchbaren Wert für a kommen sollten, sah ich mich veranlaßt, einige Übungsbeispiele mit einfachen Potenzfunktionen einzuschieben. Es erweist sich an solchen Stellen im Unterricht als günstig, mit solchen Beispielen zu beginnen, bei denen die mit DERIVE erhaltenen Ergebnisse auch ohne Computer überprüft werden können (white box - black box - Prinzip)[12].

Nach den eingeschobenen Übungen wurde das uneigentliche Integral durch ein bestimmtes approximiert. Aus dem Verlauf der symmetrischen Kurve erkannten wir, daß f(x) sehr rasch für $x \to \pm\infty$ gegen 0 zu konvergiert. Wir ersetzten daher die bisherigen Grenzen $\pm\infty$ durch ± 20.

$$\int_{-20}^{20} \frac{1}{\sqrt{(2\pi)}}\, \hat{e}^{-a x^2}\, dx$$

Dann vereinfachten wir dieses bestimmte Integral und erhielten:

$$\frac{\sqrt{2}\,\sqrt{a}\ \mathrm{ERF}\left[\dfrac{20\,|a|}{\sqrt{a}}\right]}{2\,|a|}$$

Die Schüler waren schon so an DERIVE gewöhnt, daß sie sebstverständlich annahmen, daß DERIVE mit der hier auftretenden Funktion ERF wird weiterarbeiten können. Auch wenn sie selbst keine genaue Vorstellung von dieser Funktion hatten.

Die Schüler setzten nun die Gleichung f(a)=1 an und versuchten sie zu lösen:

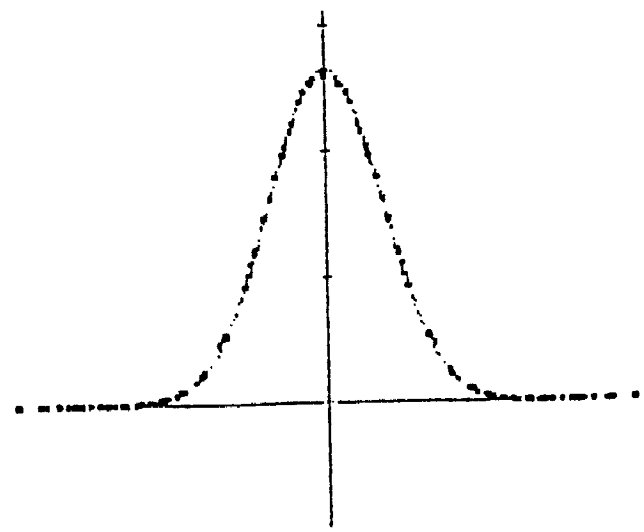
$$\frac{\sqrt{2}\,\sqrt{a}\ \mathrm{ERF}\left[\frac{20\,|a|}{\sqrt{a}}\right]}{2\,|a|} = 1$$

Da sie aber mit der Einstellung "Exact" nicht erfolgreich waren, waren sie auf meinen Vorschlag hin bereit, die Optionen auf "Approximate" zu stellen und es noch einmal mit "soLve" zu probieren. Nach einer vernünftigen Verkleinerung des angebotenen Lösungsintervall erhielten sie a = 0.499992 als Lösung. Anschließend erhöhten sie auf meinem Vorschlag die Rechengenauigkeit von 6 auf 15 Stellen, und erhielten tatsächlich a = 0.5. Die gesuchte e-Funktion war somit gefunden:

$$f(x) = \frac{1}{\sqrt{(2\,\pi)}}\,\hat{e}^{-x^2/2}$$

Die Lösung der Gleichung f(a) = 1 ist die kritische Stelle in meinem DERIVE-Konzept. Hier sind einige vorbereitende Übungen wichtig. Außerdem eine intensivere Beteuung der Schüler an den Geräten angebracht, da wirklich jeder Schüler an seinem Gerät a=0.5 erhalten sollte.

Der entscheidende Effekt tritt aber erst in dem Augenblick ein, wenn der Schüler miterlebt, wie die Kurve dieser Funktion von links kommend tatsächlich durch die vorgegebenen Punkte "hindurchläuft". Mit jedem weiteren "PLOT" entsteht die Kurve mit einer anderen Farbe aufs neue und überdeckt die alte derart, daß eine neue Mischfarbe entsteht.



Lernziel:

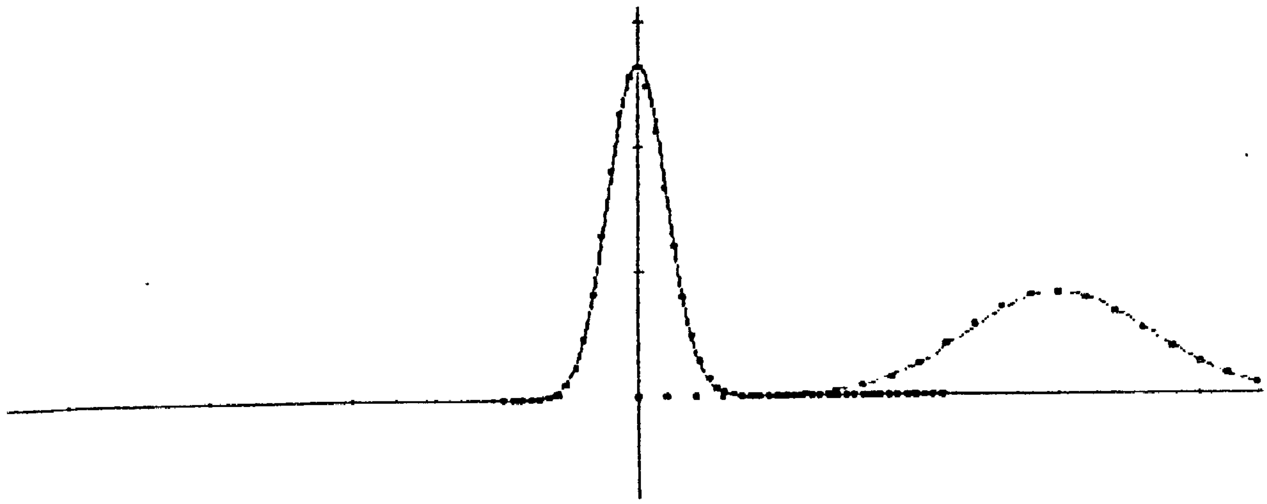
* Der Schüler soll erkennen, wie durch Experimentieren mit DERIVE die Lösung eines theoretisch für den Schüler nicht lösbaren Problems durch gezielte Approximation erraten werden kann.

5.5 Verallgemeinerung für $p \neq 0.5$

Ein weiterer Vorteil der beschriebenen Vorgangsweise mit DERIVE ist, daß der Schüler eigenständig schnell erkennen kann, daß dieselbe Funktion f(x) auch für standardisierte Binomialverteilungen $p \neq q$ (= 1-p) gilt.
Aber nicht nur die standardisierte Glockenkurve kann der Schüler jetzt zeichnen, auch die nicht standardisierte. Macht man die Transformation, die zur Standardisierung geführt hat, rückgängig, so erhält man die gewünschte e-Funktion :

$$y = \frac{1}{\sigma \sqrt{(2\ \pi)}}\ \hat{e}^{-\,0.5\ ((x\,-\,\mu)\ /\ \sigma)^2}$$

Das folgende Bild zeigt die Punkte der Binomialverteilng in standardisierter und nicht standardisierter Form für n=50 und p=0.3 mit den beiden dazugehörenden Funktionsgraphen.



Lernziel:

* Der Schüler soll an Hand der mit Hilfe von DERIVE erzeugten Bilder erkennen, daß die symmetrische Glockenform der Verteilungskurve bei großen n auch für $p \neq 0.5$ in guter Näherung gültig ist.

6. Demonstration verschiedener Lösungswege an Hand eines Beispiels

An einem Beispiel aus einem österreichischen Lehrbuch sollen fünf mögliche Lösungswege für den Schüler demonstriert werden. Bieher konnte bei diesem Beispiel - ohne Computer - nur ein Lösungsweg mit der Tabelle, mit dem ich beginnen möchte, genommen werden.

"Die Säuglingssterblichkeit, dh. die Wahrscheinlichkeit, innerhalb des ersten Lebensjahres zu sterben, betrage in einem bestimmten Land 1.8%. Wie groß ist die Wahrscheinlichkeit, daß von 1000 zufällig gewählten Säuglingen mehr als 950 und weniger als 980 Kinder den ersten Grburtstag erleben ?"[13]

Die Lösung mit Tabelle

Der Lösungsweg von Ch. Sitz [14]:

$$n=1000,\ p=0.982 \rightarrow \mu=n.p=1000.0.982 = 982$$

$$P\ (950 < X < 980) = P\ (951 \le X \le 979) =$$

$$\Phi\left(\frac{979 + 0.5 - 982}{4.204}\right) - \Phi\left(\frac{951 - 0.5 - 982}{4.204}\right) =$$

$$\Phi(-0.59463)-\Phi(-7.49236)=1-\Phi(0.59463)-(1 - \Phi(7.49236)=$$

$$1 - 0.72395 - (1 - 1) = \underline{0.27605}$$

Kritik am Lösungsweg:
Da die so durchgeführte Rechnung als vollständige Lösung - siehe Titel des Buches - angeboten wird, ist das Fehlen des bildlichen Bezugs zur Glockenkurve zu kritisieren. Wenigstens in den Musteraufgaben des Lehrbuches [15] hätte es getan werden müssen. Es besteht sonst die Gefahr, daß entgegen der Intention des Lehrplans "nur" gerechnet wird, ohne den Bezug zu den zugrundeliegenden Ideen im Auge zu behalten.

Die Lösung mit Hilfe der Binomialverteilung

$$\text{COMB}(n, k)\ p^{k}\ (1 - p)^{n - k}$$

$$\text{COMB}(1000, k)\ 0.982^{k}\ (1 - 0.982)^{1000 - k}$$

$$\sum_{k=951}^{979} \text{COMB}(1000, k)\ 0.982^{k}\ (1 - 0.982)^{1000 - k}$$

$$P(951 \le X \le 979) = \underline{0.270831}$$

Mit Hlfe von DERIVE kann man jedes Beispiel zur Binomialverteilung einfach in dieser Form rechnen. Es dauert nur manchmal etliche Minuten bis der Computer das Ergebnis ausgerechnet hat. Bestimmt man das Ergebnis anschließend mit Hilfe der Normalverteilung, so ist der Zeitaufwand beträchtlich geringer. Die Übereinstimmung der Ergebnisse nach beiden Wegen betrifft mehrere Stellen, was die Schüler nach den selbst gemachten Bilddarstellungen nicht mehr überraschte.

Die Lösung mit Hilfe des Integrals in standardisierter Form

$$\mu = n.p = 1000\ 0.982 = 982$$
$$\sigma = \sqrt{(n\ p\ (1 - p))} = \sqrt{(1000\ 0.982\ (1 - 0.982))} = 4.20428$$

$$z_1 = \frac{950.5 - 982}{4.20428} = -7.49236$$

$$z_2 = \frac{979.5 - 982}{4.20428} = -0.594632$$

$$\int_{-7.49236}^{-0.594632} \frac{1}{\sqrt{(2\ \pi)}}\ \hat{e}^{-z^2/2}\ dz = \underline{0.276044}$$

Bemerkung zum Lösungsweg:

Die zu integrierende Funktion bleibt immer gleich, die Grenzen müssen wie bei der Verwendung der $\Phi(z)$-Tabelle transformiert werden. Die Skalierung der graphischen Darstellung am Bildschirm kann bei jedem Beispiel gleich gelassen werden.

Die Lösung mit Hilfe des Integrals in nicht standardisierter Form

$$\frac{1}{\sigma\sqrt{(2\pi)}}\, ê^{-((x-\mu)/\sigma)^2/2}$$

$\mu = 982$ und $\sigma = 4.20428$

$$\frac{1}{4.20428\sqrt{(2\pi)}}\, ê^{-((x-982)/4.20428)^2/2}$$

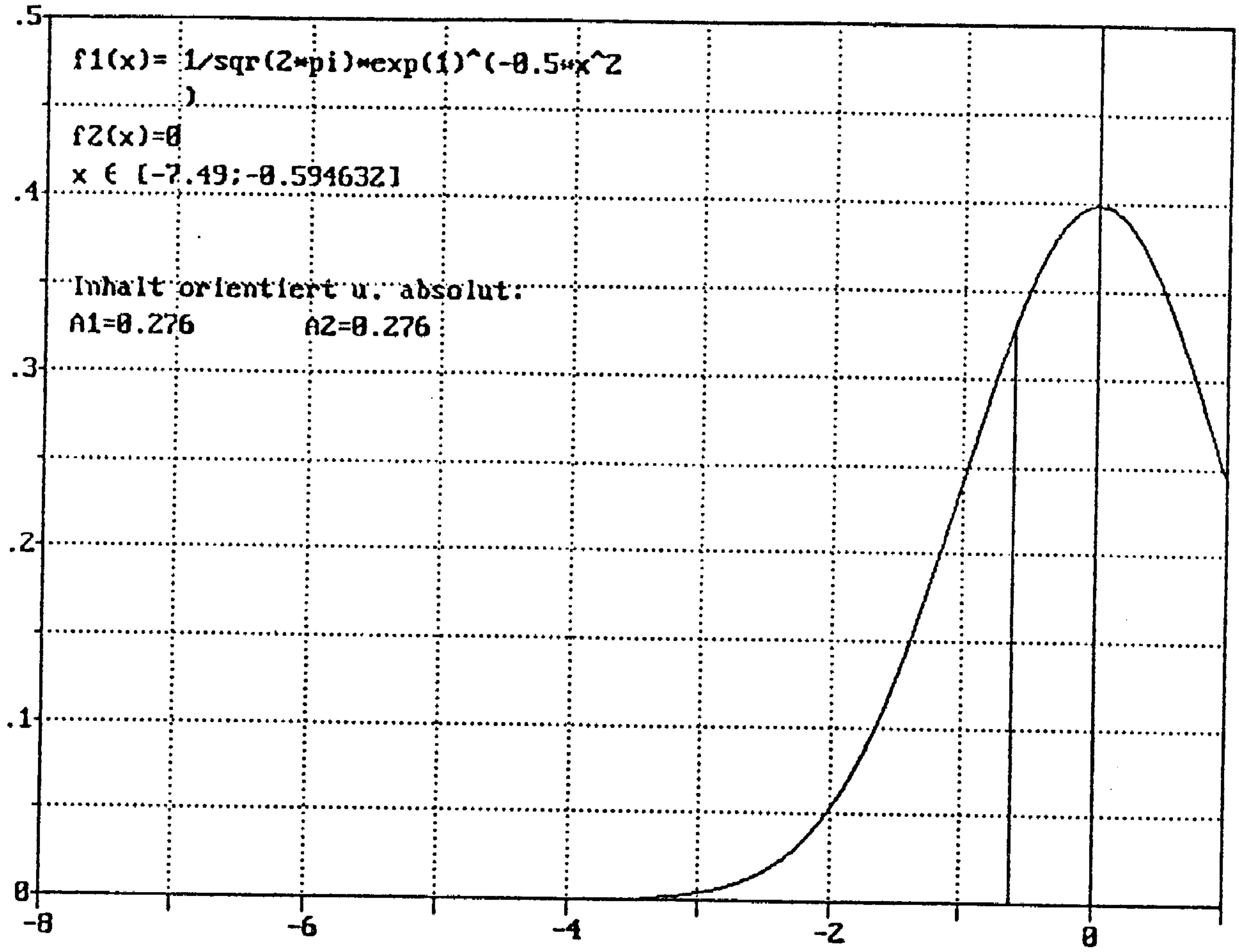
$$\int_{950.5}^{979.5} \frac{1}{4.20428\sqrt{(2\pi)}}\, ê^{-0.5((x-982)/4.20428)^2}\,dx$$

$$\frac{\mathrm{ERF}\left[\frac{393750\sqrt{2}}{105107}\right]}{2} - \frac{\mathrm{ERF}\left[\frac{31250\sqrt{2}}{105107}\right]}{2} = \underline{0.276045}$$

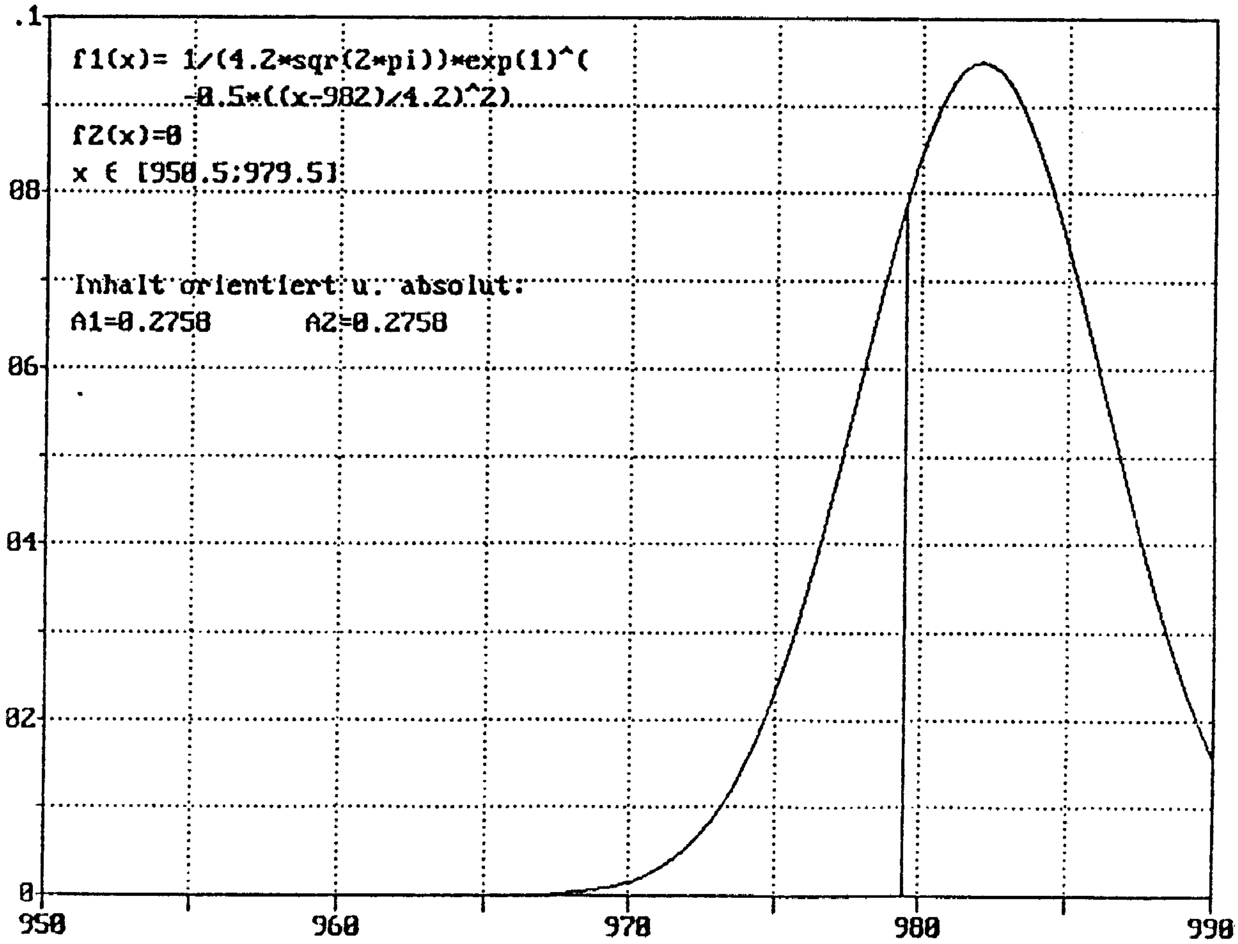
Bemerkung zm Lösungsweg:
In diesem Fall bleiben die Grenzen gleich, was das Verständnis deutlich verbessert, dafür ist die zu integrierende Funktion komplizierter. Vor allem bereiten die vielen, einzugebenden runden Klammern beachtliche Probleme, weshalb dieser Weg fast nur von den computererfahrenen Schülern gewählt wurde. Die Skalierung der graphischen Darstellung am Bildschirm muß bei jedem Beispiel neu überlegt werden. Ist sie aber günstig gewählt, so ist die Zeichnung leichter zu verstehen als der standardisierten Form.

Die Lösung mit MATHEASS

Mit MATHEASS läßt sich nicht nur fast jedes bestimmte Integral rasch berechnen. Man erhält außerdem ein anschauliches Bild am Computer dazu. Ich habe daher MATHEASS häufig im Klassenraum am Demonstrationscomputer verwendet.



f1(x)= 1/sqr(2*pi)*exp(1)^(-0.5*x^2
)
f2(x)=0
x ∈ [-7.49;-0.594632]
Inhalt orientiert u. absolut:
A1=0.276
A2=0.276
.5
.4
.3
.2
.1
0
-8
-6
-4
-2
0



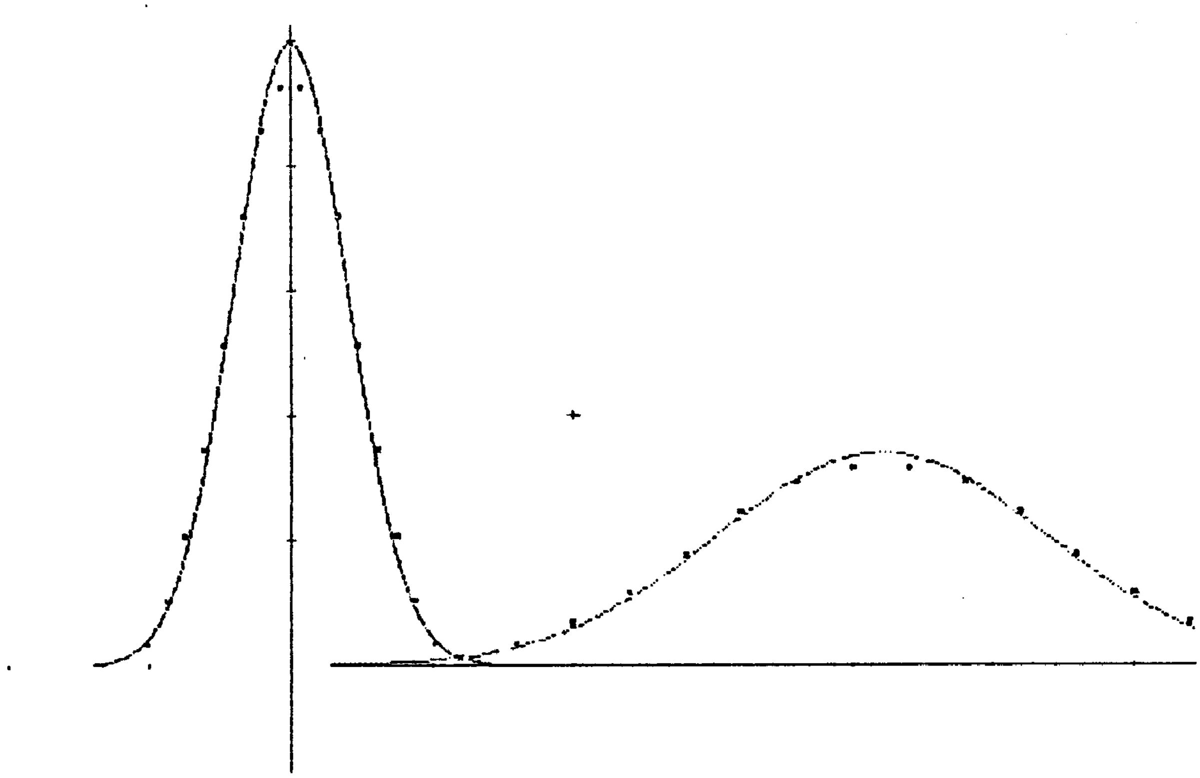
f1(x)= 1/(4.2*sqr(2*pi))*exp(1)^(
-0.5*((x-982)/4.2)^2)
f2(x)=0
x ∈ [958.5;979.5]
Inhalt orientiert u. absolut:
A1=0.2758
A2=0.2758
.1
08
06
04
02
0
950
960
970
980
990

Bemerkung zu Matheass:
Es ist wirklich nicht einfach die e-Funktion mit μ und σ richtig mit allen runden Klammern im ersten Versuch einzugeben. Gelingt es aber und wählt man die Skalen günstig, so erhält man ein sehr anschauliches Bild. Matheass kann nur dann zur Lösung von Aufgaben zur Normalverteilung herangezogen werden, wenn der Flächeninhalt als Maß für die Wahrscheinlichkeit zu berechnen ist. Sind aber die Intervallgrenzen oder μ oder σ gesucht, so hat man DERIVE zu verwenden.

7. Ein weiterer Zugang zur Normalverteilung

Ich habe bisher ein Unterrichtskonzept dargestellt, in dem die Verteilungsfunktion der Normalverteilung über die Binomialverteilung gefunden wurde. Um die Eigenständigkeit der Normalverteilung im Unterricht herauszuarbeiten, ist es aber günstig, den Schülern noch einen anderen Zugang zu zeigen.
Ein solcher wird im Buch "Wahrscheinlichkeitsrechnung und Statistik" von Brenner-Lesky-Vogel im Kapitel "Normalverteilung" [16] vorgeschlagen. Es wird mit drei Würfel-Beispielen begonnen. Die Augenzahlen eines Würfels genügen einer diskreten Gleichverteilung, die Augensummen zweier Würfel einer diskreten Dreiecksverteilung und die Augensummen dreier Würfel nähern sich einer glockenförmigen Verteilung.
Die Vorteile dieses Einstiegs liegen einerseits in der Möglichkeit zu eigenem Experimentieren, andererseits darin, daß bei der Behandlung von Wahrscheinlichkeitsverteilungen in der 7. Klasse diese Überlegungen bereits vorbereitet worden sind. Daher habe ich diesen Zugang in mein computerunterstütztes Konzept eingebaut.
Ich habe also mit den Schülern die Verteilung der Augensummen dreier Würfel mit Hilfe des Computers untersucht. Die Werte der Verteilung und den Erwartungswert μ riefen wir sofort durch das Stochastik-Programm von Stormer ab. Die Varianz σ aber ließ ich die Schüler selbst berechnen, damit ihnen bewußt wird, daß es sich dieses Mal um keine Binomialverteilung handelt.

Dann zeichneten alle mit Hilfe von DERIVE den Punktgraphen. Wir stellten die Noralverteilungsfunktion mit $\mu=10.5$ und $\sigma\approx2.958$ auf und mit "PLOT PLOT" erhielten wir das folgende Bild:



In Anlehnung an Brenner-Lesky-Vogel [17] formulierte ich den zentralen Grenzwertsatz:

Sind die voneinander (stochastisch) unabhängigen Zufallsvariablen X_i $(i=1,2,\dots,n)$ mit μ und σ identisch verteilt, dann nähert sich die Wahrscheinlichkeitsverteilung der Zufallsvariablen $X = \Sigma X_i$ mit wachsendem n immer mehr der Normalverteilung.

Dieser Satz bleibt unter geringen einschränkenden Voraussetzungen, die in der Praxis fast immer erfüllt sind, für nicht identisch verteilte Zufallsvariable gültig.

Anmerkungen und Literaturhinweise:

[1] BGBl. Nr. 275/1970.

[2] BGBl. Nr. 114/1978.

[3] BGBl. Nr. 327/1989.

[4] Novak, Mathematik Oberstufe 3. Reniets Verlag, Wien 1991, S. 59f.

[5] ebd. S. 64.

[6] Reichel-Müller-Hanisch, Lehrbuch der Mathematik 8. Verlag Hölder-Pichler-Tempsky, Wien 1992, S. 111.

[7] ebd. S. 127-132.

[8] Szirucsek-Dinauer-Unfried-Schatzl, Mathematik 8, Verlag Hölder-Pichler-Tempsky, Wien 1992, S. 132f.

[9] BÜrger-Fischer-Malle, Mathematik Oberstufe 4, Verlag Hölder-Pichler-Tempsky, Wien 1992, S. 107.

[10] ebd. S. 110.

[11] Malle G., Wahrscheinlichkeitsrechnung und Statistik in der 7. und 8. Klasse, in: Mathematik AHS-Oberstufe Kommentar, Österreichischer Bundesverlag, Wien 1991, S. 168.

[12] Buchberger B., Should Students Learn Integration Rules ?, RISC-LINZ,Tech.Rep.89-7.0, Johannes Kepler Universität Linz, 1989.

[13] in [6], Beispiel 432c, S. 124.

[14] Sitz, Ch., Vollständige Lösungen zu Wahrscheinlichkeitsrechnung und Statistik 7. und 8. Klasse, Verlag Hölder-Pichler-Tempsky, Wien 1992, 432c, S. 68.

[15] in [6], Musteraufgaben M - R, S. 129-137.

[16] Brenner/Lesky/Vogel, Wahrscheinlichkeitsrechnung und Statistik, Klett Verlag, Stuttgart 1980.

[17] ebd. Seit 113.