Günter HANISCH

Von der Binomialverteilung zur Normalverteilung ein Unterrichtsvorschlag*

1. Einleitung

In meiner letzten Arbeit in dieser Reihe (Hanisch, 1984a) brachte ich folgendes Beispiel:

Beispiel 1:
Im Österreichischen Nationalrat sind von 183 Abgeordneten 17 weiblich. Wenn wir annehmen, daß eine Frau die gleiche Chance wie ein Mann hat, in den Nationalrat entsendet zu werden, wie groß ist dann die Wahrscheinlichkeit, daß 17 oder noch weniger Frauen in den Nationalrat entsendet werden.
Bei der Lösung schrieb ich u.a. etwa so: Wegen des zu großen Rechenaufwandes wird die Binomialverteilung durch die Normalverteilung approximiert.

In dieser Arbeit möchte ich nun einen Weg zeigen, wie man diese Approximation den Schülern näher bringen kann, denn m.E. können sich bei der Einführung der Normalverteilung im Unterricht drei Schwierigkeiten ergeben:

1) Die Normalverteilung ist eine kontinuierliche (stetige) Verteilung. Im Unterricht wurden aber i.a. nur diskrete Verteilungen besprochen.
2) Die Funktionsgleichung der Dichtefunktion wird mitgeteilt, nicht bewiesen oder hergeleitet.
3) Das Integral über die Dichtefunktion ist nicht geschlossen darstellbar.

Allerdings wird in dieser Arbeit fast nicht darauf eingegangen, ob - und wenn ja - warum die Normalverteilung überhaupt besprochen werden soll. Läßt man sie weg, kann man m.E. auf die kontinuierlichen Verteilungen überhaupt verzichten, denn Tests und Konfidenzintervalle können auch ohne jene durchgeführt werden. Allerdings bilden die kontinuierlichen Verteilungen

* Die Grundlage der Idee zu dieser Arbeit ist durch die Vorlesung über Schulmathematik V von Prof. REICHEL und insbesondere durch unsere Arbeit an dem Werk "Didaktik der Stochastik" (Arbeitstitel) entstanden, das demnächst im HPT erscheinen wird.

ein ideales Bindeglied zur Analysis, und die am häufigsten angewandten Verfahren der schließenden Statistik können ohne Normalverteilung oder deren Abkömmlinge nicht auskommen.

2. Relative Häufigkeitsdichte

M.E. sind es folgende Probleme im MU, die kontinuierliche Verteilungen im Gegensatz zu diskreten aufweisen:

1. Die Wahrscheinlichkeit eines Einzelereignisses ist Null.
2. Die Wahrscheinlichkeit eines Intervalls wird als Fläche unter einer Dichtefunktion, die anstatt der Wahrscheinlichkeitsfunktion auftritt, berechnet.

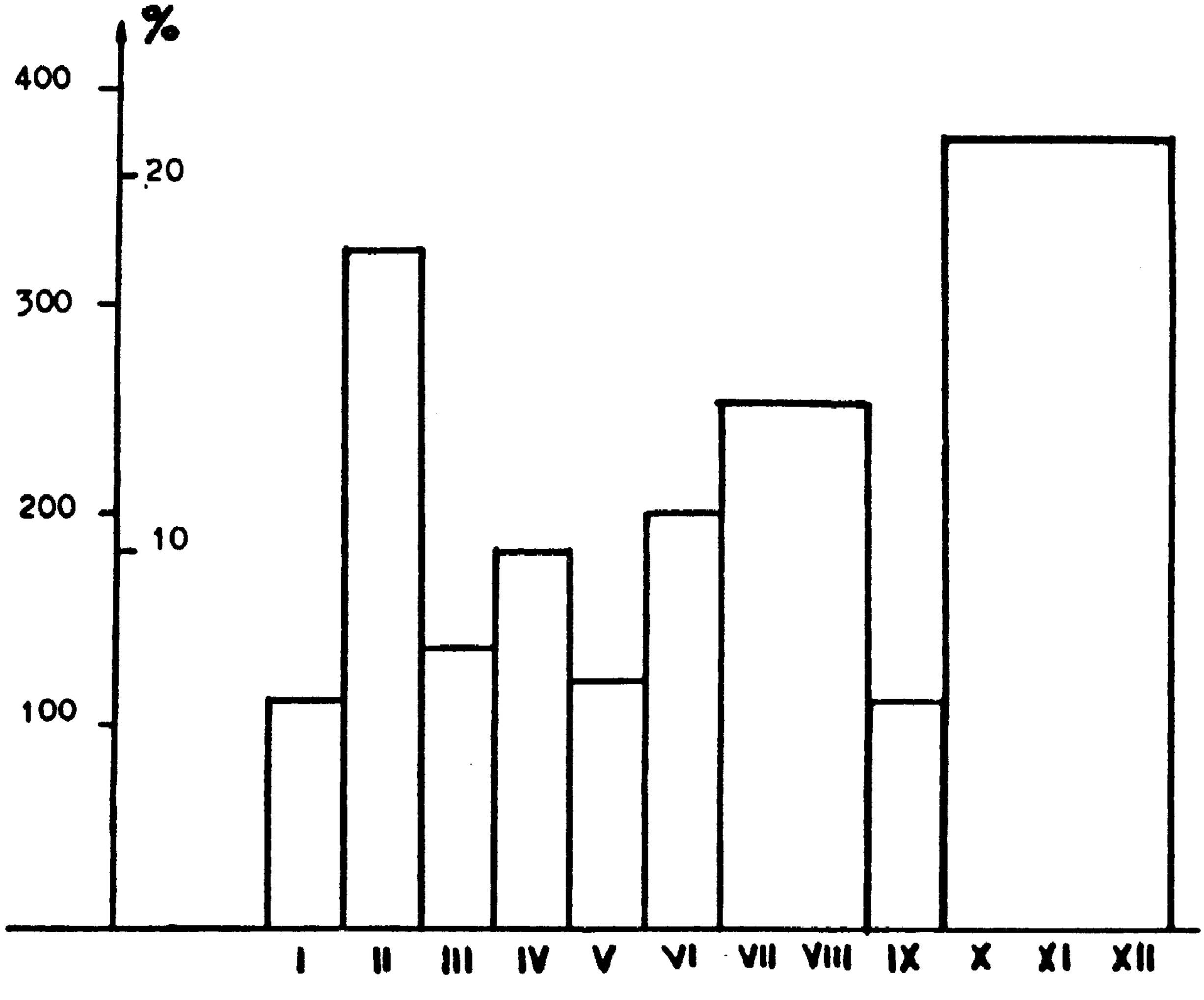
Hat man nun vor im Unterricht auch kontinuierliche Verteilungen zu besprechen, so sollten bereits bei der beschreibenden Statistik die Grundlagen dafür gelegt werden:

Beispiel 2: Franz erhielt im vergangenen Jahr folgende Geldzuwendungen (Taschengeld, Geschenke von den Großeltern, der Tante etc.).

Monat	a_n (Betrag in S)	h_n in %
Jänner	108.--	6,0
Februar (Geburtstag)	324.--	18,0
März	135.--	7,5
April (Ostern)	180.--	10,0
Mai	117.--	6,5
Juni (Schulschluß)	198.--	11,0
Juli, August	252.--	14,0
September	108.--	6,0
Oktober, November, Dezember	378.--	21,0
Summe	1800.--	100,0

Dabei bezeichnet a_n die Häufigkeit und $h_n = \frac{a_n}{n}$ die relative Häufigkeit.

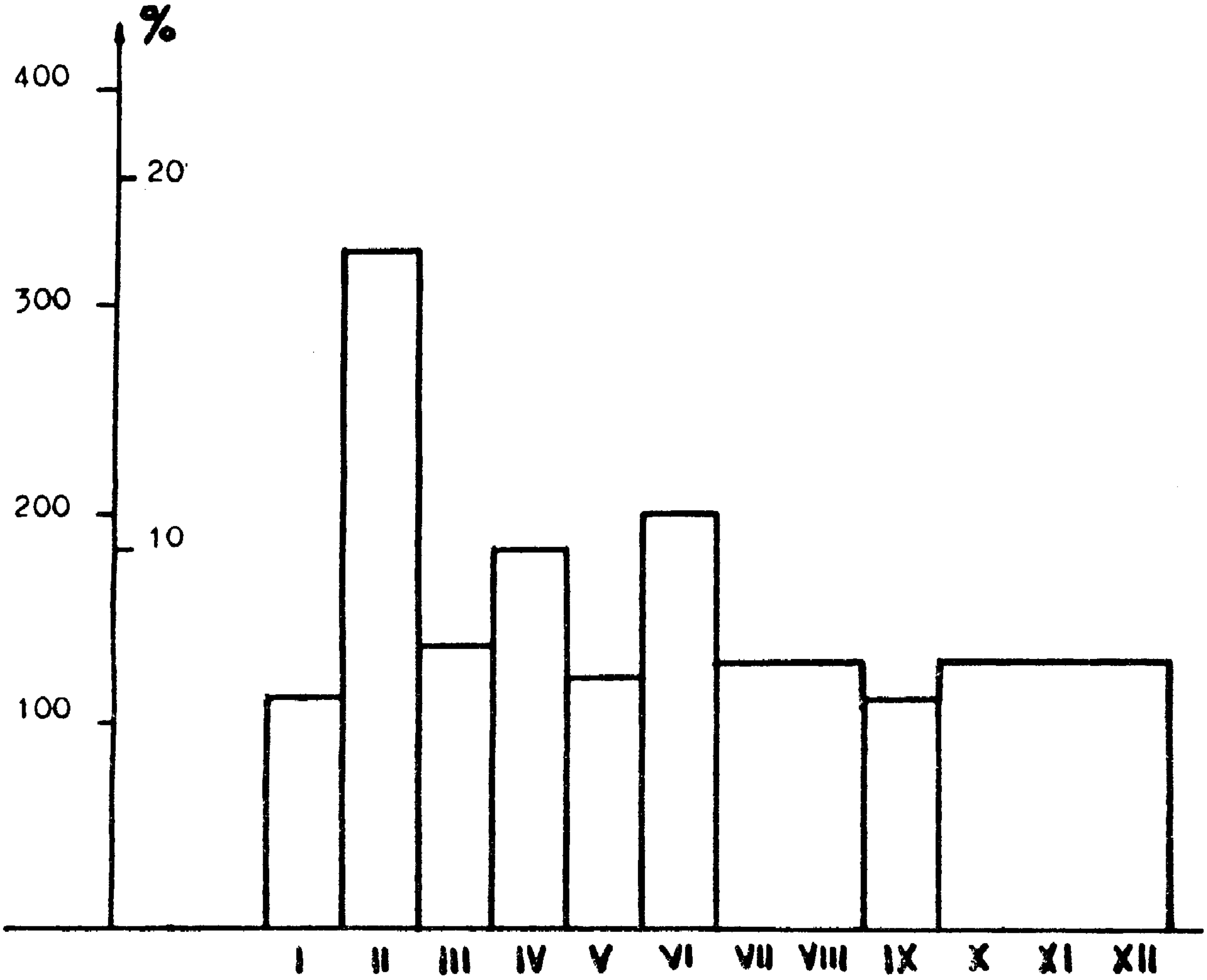
Stellen wir seine Einkünfte in einem Balkendiagramm dar, so ergibt sich die Frage, wie wir die Einkünfte des Juli und des August und die von Oktober bis Dezember eintragen sollen:
Tragen wir nämlich die Einkünfte mit ihrem totalen Betrag auf, so ergibt sich ein völlig falscher Eindruck, weil man unwillkürlich nicht die Höhe, sondern den Flächeninhalt als Maß für die Größe des Geldbetrages nimmt.



Um daher nicht ein subjektiv falsches Bild zu erhalten, ist es sinnvoll, die Ordinate (Häufigkeit bzw. rel. Häufigkeit) so abzuändern, daß nur der auf den entsprechenden Monat entfallende Anteil aufgetragen wird. Wir tragen daher nicht die Häufigkeit, sondern $\frac{\text{Häufigkeit}}{\text{Klassenbreite}}$ = Häufigkeitsdichte $\overline{a_n}$

bzw. $\frac{\text{relative Häufigkeit}}{\text{Klassenbreite}}$ = relative Häufigkeitsdichte $\overline{h_n}$ *

auf.

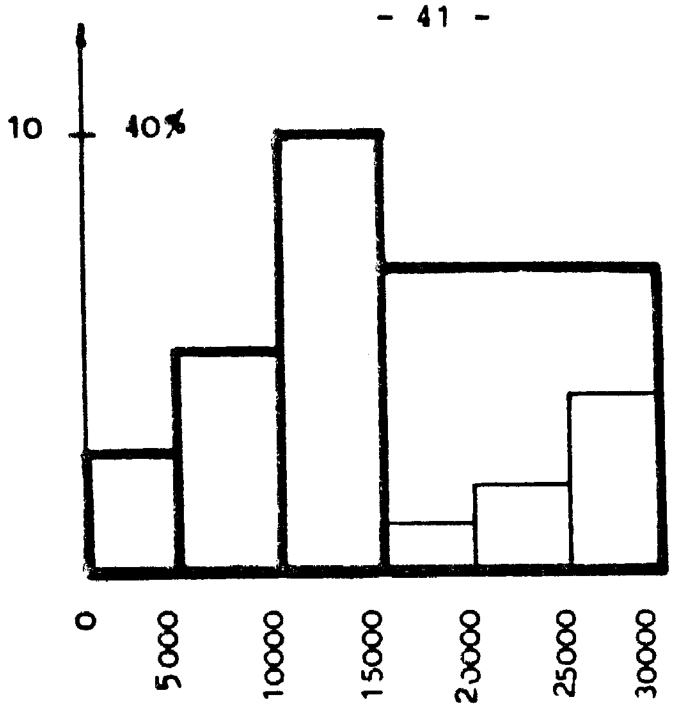


Diese Graphik gibt einen "richtigen" Eindruck von der Jahresverteilung wieder, denn als Maß tritt der

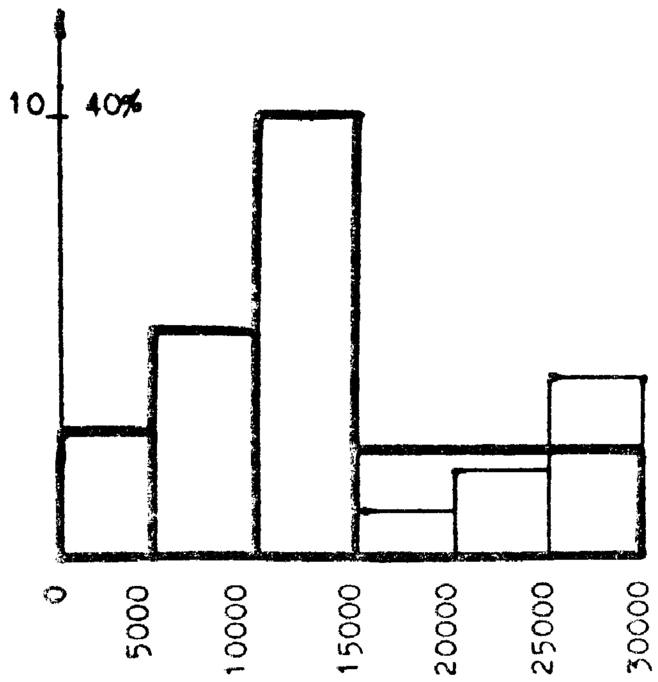
Flächeninhalt = relative Häufigkeitsdichte · Klassenbreite *
= relative Häufigkeit

auf.

* Diese beiden Begriffe scheinen zwar in der Literatur nicht auf; m.E. aber ist es zweckmäßig sie einzuführen, zumal dadurch im Sinne des Spiralprinzips bereits hier auf den Dichtebegriff hingewiesen wird.



Das zweite Balkendiagramm suggeriert so wie vorhin den Eindruck, daß das Einkommensniveau viel höher liegt. Es ist daher wieder besser, bei ungleichen Klassenbreiten als Maß den Flächeninhalt des Balkens zu wählen, also die Häufigkeitsdichte bzw. die relative Häufigkeitsdichte aufzutragen:



Bemerkung: Berechnet man die Summe der Flächeninhalte aller Balken, so erhält man:

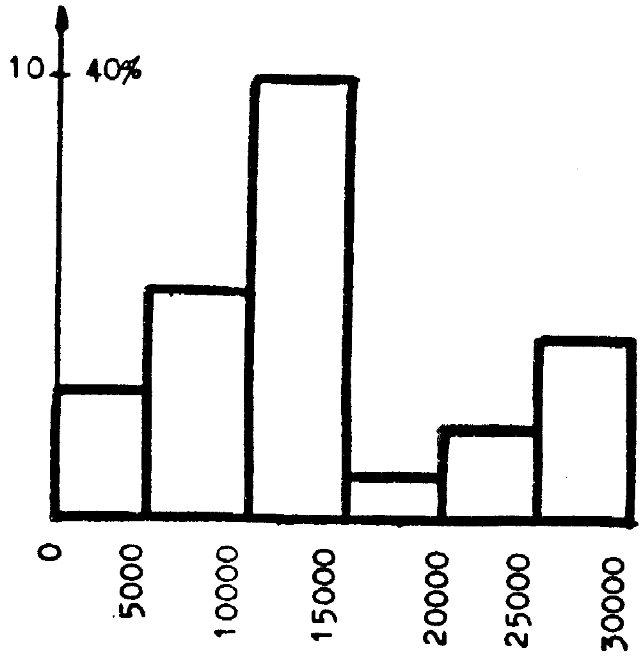
$$\Sigma\, \bar{h}_i \cdot d_i = \Sigma\, \frac{h_i}{d_i} \cdot d_i = \Sigma h_i = 1$$

Die Summe der Flächeninhalte aller Balken ist also 1.

<u>Beispiel 3:</u> Das Gehaltsniveau einer Firma sei durch folgende Tabelle charakterisiert:

Gehalt	Anzahl
0- 5000	3
5001-10000	5
10001-15000	10
15001-20000	1
20001-25000	2
25001-30000	4
	25

Balkendiagramm:



Fassen wir hingegen die letzten drei Klassen zu einer zusammen, so ergibt sich folgendes Bild:

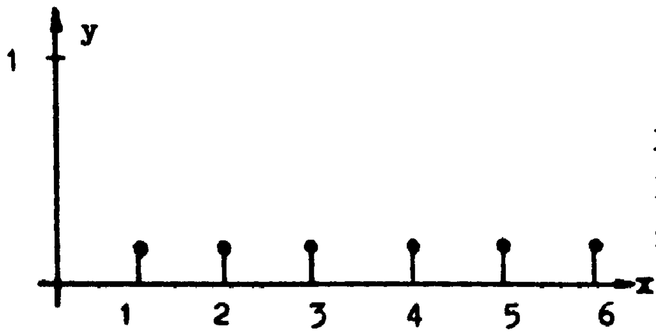
3. Verteilungs- und Dichtefunktion

Als nächstes wollen wir uns dem Begriff "Zufällige Variable" zuwenden.

Eine Größe X, die - etwa als Ergebnis eines beliebig wiederholbaren Experiments - verschiedene reelle Zahlenwerte - annehmen kann, bezeichnen wir dann als zufällige Variable, wenn wir den Wert dieser Variablen (prinzipiell) nicht vorhersagen können. Wir unterscheiden diskrete und kontinuierliche Variable, je nachdem, ob X nur endlich (oder abzählbar) viele Werte annehmen kann oder alle Werte eines bestimmten Intervalls der Zahlengeraden.

Bemerkung: Wichtig erscheint es m.E., die Schüler genau darauf hinzuweisen, daß X die Zufallsvariable ist, die Werte, die sie annimmt, aber mit x bezeichnet werden.

Beispiel 4: Beim Werfen eines idealen Würfels sei X die Augenzahl. Hier kann die zufällige Variable leicht so dargestellt werden, daß zu jedem Wert x, den X annehmen kann, die entsprechende Wahrscheinlichkeit angegeben wird: $y(x) = P(X=x)$.

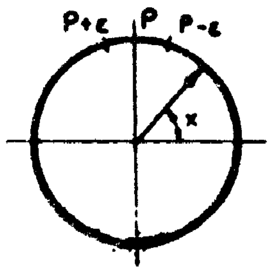


Diese so dargestellte (diskrete) Funktion heißt Wahrscheinlichkeitsfunktion.

Die Wahrscheinlichkeitsfunktion hat zwei Nachteile:

(1) Es handelt sich - graphisch gesehen - um lauter Einzelpunkte, es liegt also keine "Kurve" vor, diese - für Schüler wohl-typische Form einer Funktion.

(2) Für kontinuierliche Zufallsvariable (Winkel beim Glücksrad oder Durchmesser eines Werstückes) ist diese Darstellung wertlos, da u.a. für einen festen x-Wert $P(X=x)=0$ gilt. Diesen für kontinuierliche Zufallsvariable typischen Fall zeigt das folgende Beispiel ganz deutlich:



Beispiel 5: Glücksrad
Der Winkel $X=a$, bei welchem der Zeiger stehen bleibt, ist eine kontinuierliche Zufallsvariable. Die Wahrscheinlichkeit, daß der Zeiger genau bei $X=90^{\circ}$ stehen bleibt, muß Null gesetzt werden, da ja
$P(\text{Zeiger bleibt im Intervall } (P-\varepsilon, P+\varepsilon) \text{ stehen}) = \frac{2\varepsilon}{2r\pi} \to 0$ für $\varepsilon \to 0$. Und das gilt analog für jeden anderen Punkt des Umfangs. Trotzdem kann der Zeiger natürlich genau bei 90° stehen bleiben.

Die Wahrscheinlichkeitsfunktion, wie wir sie oben definiert haben, wäre also die konstante Nullfunktion: $y(x) = 0 \forall x$. Zur graphischen Darstellung verwendet man daher bei kontinuierlichen Zufallsvariablen - wie wir gleich sehen werden - andere Funktionen (nämlich die Verteilungs- oder auch - wenn möglich - die Dichtefunktion).

Beispiel 6: Auf einer Drehbank werden gewisse Werkstücke hergestellt, deren Durchmesser 200 mm betragen soll. Der tatsächliche Durchmesser ist eine Zufallsgröße, es werden Durchmesser vorkommen, die größer oder kleiner als 200 mm sind.

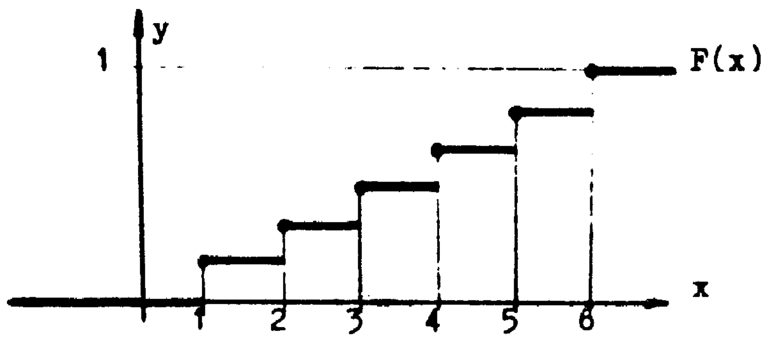
Auch hier wird die Wahrscheinlichkeit, daß der Werkstückdurchmesser im Intervall $[l-\varepsilon, l+\varepsilon]$ liegt, mit sinkender Intervallänge abnehmen und daher bei $\varepsilon \to 0$ auch gegen 0 gehen. Somit gilt für jedes feste l:

$$P(X=l)=0 .$$

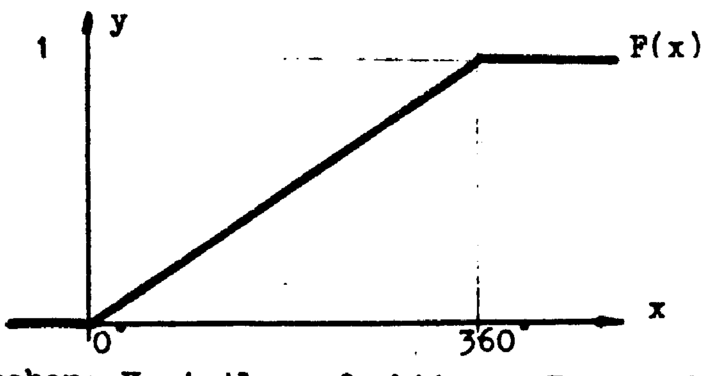
Aus diesem Grund wählen wir für kontinuierliche zufällige Variable eine andere Darstellung: die sogenannte <u>Verteilungsfunktion.</u> Wir definieren sie so, daß sie auch gleich für diskrete Zufallsvariable anwendbar ist. Verteilungsfunktionen weisen keinen der sogenannten "Nachteile" mehr auf, sie bilden eine <u>einheitliche Darstellungsform</u> für diskrete und kontinuierliche Zufallsvariable, sie sind "universeller" anwendbar.

Die Idee ist einfach: Man gibt statt $y(x) = P(X=x)$ den Wert $y=F(x)=P(X\leq x)$ an. Diese definierte Funktion F heißt dann <u>Verteilungsfunktion</u> der Zufallsvariablen X.
Wir studieren sie gleich an unseren Beispielen:

<u>Beispiel 4</u> (Fortsetzung): Würfelwurf



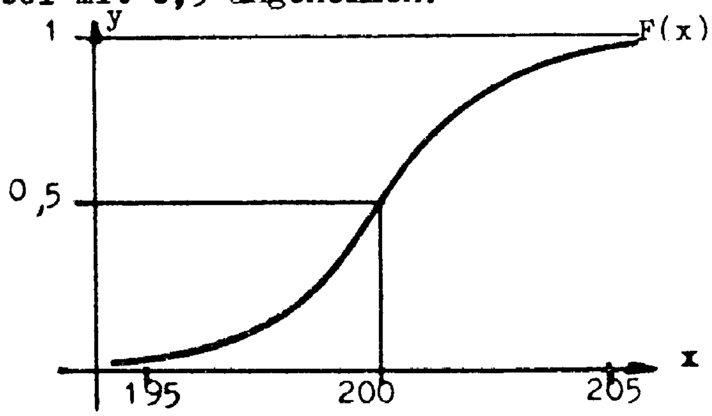
<u>Beispiel 5</u> (Fortsetzung): Glücksrad



Wir sehen: Verteilungsfunktionen F sind für <u>jedes</u> <u>$x\in\mathbb{R}$</u> erklärt (für $x < 0$ setzt man bei unseren Beispielen einfach $F(x)=0$).

<u>Beispiel 6</u> (Fortsetzung): Werkstückbeispiel

Die Verteilungsfunktion könnte etwa so aussehen, $P(X \leq 200)$ ist dabei mit 0,5 angenommen.



Daß wir-bei Angabe der Verteilungsfunktion - auf die Wahrscheinlichkeit <u>des</u> <u>Einzelereignisses</u> (also $X=x$ für festes x) verzichten, bringt de facto keinen Informationsverlust. Diese Wahrscheinlichkeit interessiert uns ja in der Praxis gar nicht. Was heißt denn, der Durchmesser eines Werkstücks beträgt 199,8 mm? Das ist doch eine Frage der Meßgenauigkeit. In Wirklichkeit interessieren wir uns etwa dafür, daß der Durchmesser in das Intervall (19,974; 19,985] fällt. Das aber können wir aus der Verteilungsfunktion leicht berechnen. Es ist

$$P(19{,}974 < X \leq 19{,}985) = P(X \leq 19{,}985) - P(X \leq 19{,}974) =$$
$$= F(19{,}985) - F(19{,}974)$$ und allgemein gilt:
$$P(a < X \leq b) = P(X \leq a) = F(b) - F(a).$$

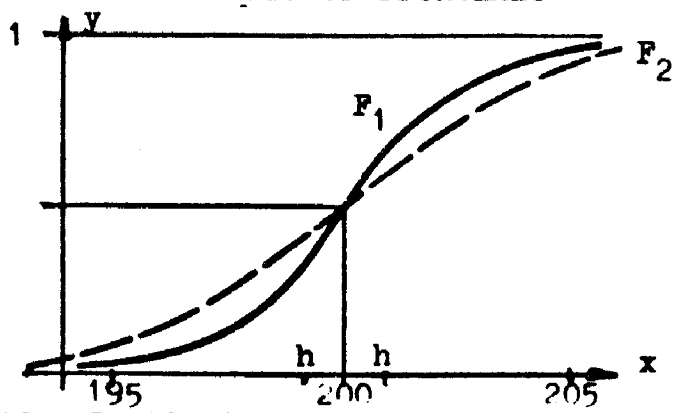
Ob man dabei offene, halboffene oder geschlossene Intervalle nimmt, spielt keine Rolle, da für jedes a $P(X=a)=0$.
Es gilt daher auch:
$$P(a \leq X < b) = F(b) - F(a) ,$$
aber das nur "nebenbei".

Es ist unmittelbar klar, daß Verteilungsfunktionen <u>monoton wachsend</u> sind und $\lim\limits_{x \to \infty} F(x) = \infty$, $\lim\limits_{x \to \infty} F(x) = 0$. Sie müssen aber nicht überall stetig sein, wie wir gesehen haben.

<u>Bemerkung:</u> Verteilungsfunktionen können also Sprungstellen haben, z.B. bei diskreten Zufallsvariablen (Beispiel 3). Prinzipiell können aber auch Verteilungsfunktionen von kontinuierlichen Variablen Sprungstellen haben; denken Sie sich als Beispiel eine kontinuierliche Größe, die Sie mit einer diskreten überlagern.

In der Praxis sind Verteilungsfunktionen (bei ihren Stetigkeitspunkten) meist sogar <u>differenzierbar.</u> Der Differentialquotient $F'(x)$ gibt uns dann wichtige Informationen über die Zufallsvariable X ("Dichtefunktion"). Wir wollen dies nun genauer behandeln, d.h. wir interessieren uns nun dafür, was die Steilheit der Kurve bedeutet. Im Unterricht kann man an dieser Stelle bereits "experimentell", anhand verschiedener Kurvensteigungen etwa bei $x = 200$, die Bedeutung des Anstieges entdecken lassen:

<u>Beispiel 6</u> (Fortsetzung):
Betrachten wir die Verteilungsfunktionen der Durchmesser von Werkstücken zweier verschiedener Drehbänke

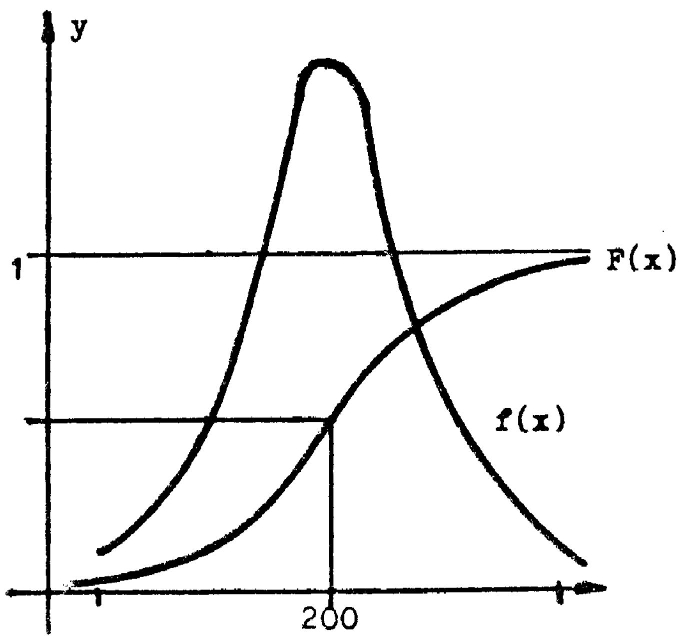


Welche der beiden Drehbänke arbeitet genauer?
<u>Antwort:</u> Es ist diejenige Maschine, die Werkstücke erzeugt, bei denen die Verteilungsfunktion der Durchmesser bei 200 mm steiler ansteigt. Denn bei dieser "konzentrieren" sich die Längen um den Wert 200 stärker, denn hier ist, selbst für kleine h, $F(x+h)$ offenbar viel größer als $F(x)$. Es ist also zu erwarten, daß

Werkstücke mit einem Durchmesser von "nahezu" 200 mm relativ häufig produziert werden. Eine starke (schwache) Steigung bedeutet also, daß viele (wenige) Werkstücke in kleinen Intervallen um l = 200 mm liegen. Die Steigung $F'(x)$ der Verteilungsfunktion F bei x_0 - vorausgesetzt F ist dort differenzierbar - gibt also gewissermaßen an, wie "dicht" die tatsächlichen Werte um den Wert x_0 liegen.
Es ist also sinnvoll, die Funktion $f = F'$ als Dichtefunktion der Zufallsvariablen zu bezeichnen.
Maschine 1 ist besser als Maschine 2, weil beim Sollwert $x_0 = 200$ die Dichte $f_1(x_0) = F_1'(x_0)$ größer ist als $f_2(x_0) = F_2'(x_0)$.

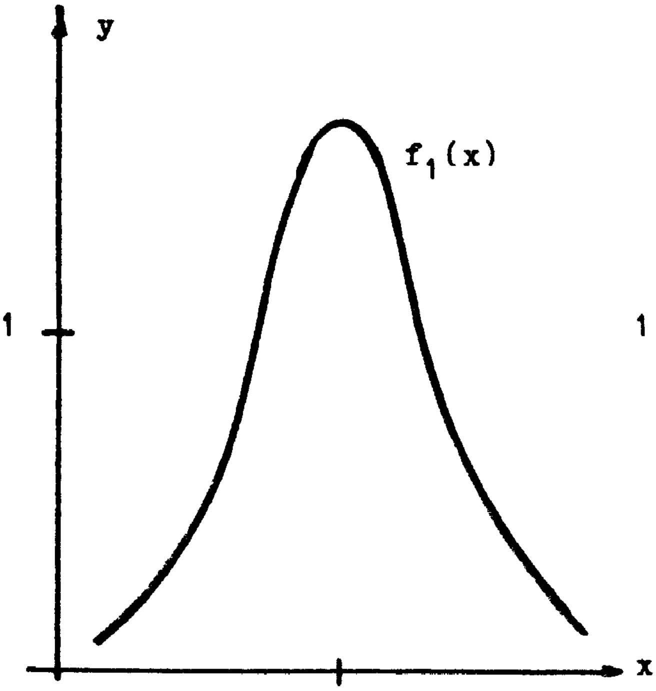
Sehen wir uns die Sache noch an einer anderen Stelle, etwa bei $x = 195$ mm, an: Bei $x = 195$ ist die Tangentensteigung bei beiden Maschinen nicht sehr groß, die Funktion $F(x)$ wächst nur geringfügig, d.h. für kleine h ist $F(x+h)$ nicht sehr viel größer als $F(x)$. Werkstücke mit Durchmesser zwischen x und x+h werden also nur relativ selten produziert werden.

Zusammenfassend können wir also sagen: der Anstieg der Tangente an F bei x $F'(x) = f(x)$, also die Änderungsrate von F, gibt die zu erwartende "Konzentration" der zu erwartenden tatsächlichen Werkstücklängen an, also die Dichte der Zufallsvariablen X. f heißt deshalb die Dichtefunktion von f .*
Man verdeutliche sich genau die Bedeutung folgender Skizze:

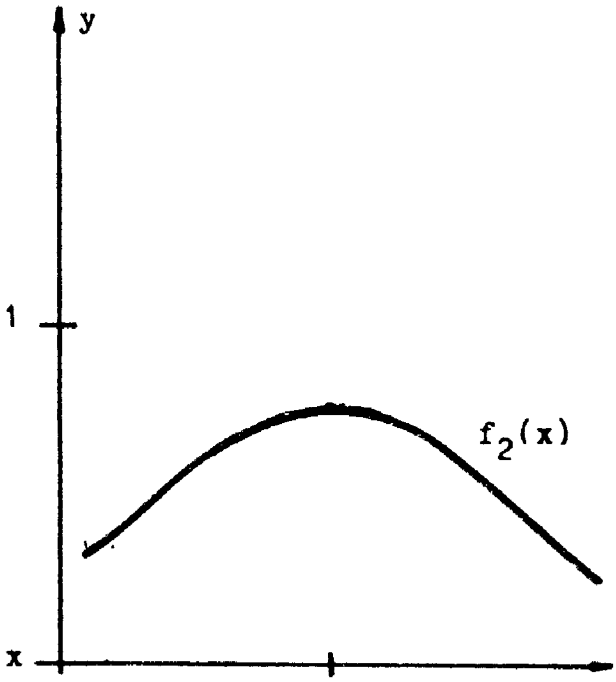


* Wird später noch exaktifiziert.

Die Graphen der beiden Wahrscheinlichkeitsdichten $f_1(x)$ und $f_2(x)$ zeigen noch deutlicher den Unterschied zwischen der Genauigkeit der beiden Maschinen.
(Die beiden Graphen können durch graphisches Differenzieren aus $F_1(x)$ und $F_2(x)$ gewonnen werden.):



schmäler, höher



breiter, niedriger

<u>Bemerkung 1</u>: Zu beachten ist, daß $f(x)$ unter Umständen Werte größer als 1 annehmen kann (Unterschied zur Wahrscheinlichkeitsfunktion), die Verteilungsfunktion $F(x)$ kann ja ohne weiteres einen Anstieg größer als 45° haben. Da $F(x)$ aber monoton nicht fallend ist, kann $f(x)$ keine negativen Werte annehmen.
<u>Bemerkung 2:</u> Zum Namen "Dichte": Denken wir uns - heuristisch gesehen - die x-Achse mit "Masse" belegt. Links von jedem festen x soll die Masse $y = F(x)$ liegen. $\frac{F(x+h)-F(x)}{h}$ ist dann die "Massendichte" im Intervall $[x;x+h]$ und $F'(x) = \lim_{h \to 0} \frac{F(x+h)-F(x)}{h}$ ist die

"Dichte" im Punkt x (wenn $F'(x)$ existiert).

Hinweis auf ein "Paradoxon": Die Masse in einem Punkt ist natürlich Null. Trotzdem ist $F'(x_0) = f(x_0)$ ein gutes Maß für die Massendichte im Punkt a. Das ist genau wie bei der Definition der "Augenblicksgeschwindigkeit" eines sich nach der Gleichung $s = F(t)$ (s... Weg; t... Zeit) bewegten Massenpunktes. Das schon bei den Griechen (Sophisten)bekannte Paradoxon ("Es gibt keinen fliegenden Pfeil,weil der Pfeil (Foto!) in jedem festgewählten Augenblick t_0 "ruht" und die Aneinanderreihung ruhender Pfeile kann ja keine Ortsveränderung ergeben".) "löst" sich dadurch, daß es eben "bloß" sinnvoll ist, dem Massenpunkt zum Zeitpunkt t_0 den Wert $F'(t_0)$ als Augenblicksgeschwindigkeit - also definitionsgemäß - zuzuschreiben.

Bemerkung 3: Die Verflechtung der Gebiete Differential- und Integralrechnung einerseits und Wahrscheinlichkeitstheorie andererseits möge man deutlich machen! Die bezuglose (oder den Schülern so erscheinende) Aneinanderreihung der einzelnen Stoffgebiete gehört zu den schlimmsten Erscheinungen des MU.

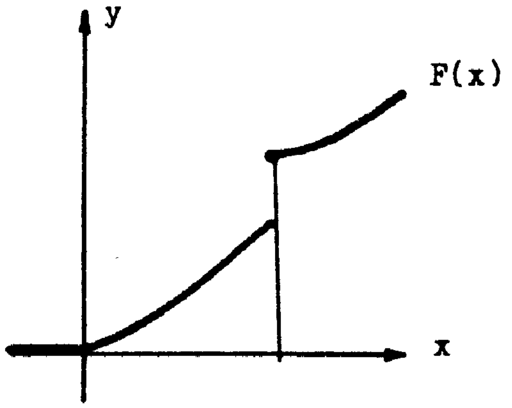
Zusammenhang zwischen Verteilungs- und Dichtefunktion:
Wir wissen: Wenn $f(x) = F'(x)$ gilt, dann ist

$$\int_{-\infty}^{x} f(t)dt = F(t) \Big|_{-\infty}^{x} = F(x) - \lim_{t\to-\infty} F(t) = F(x) - 0 = F(x)$$

Wir definieren daher(Exaktifizierung, aber m.E. nicht für den MU): Die Zufallsvariable X mit der Verteilungsfunktion F hat eine Dichtefunktion, wenn es eine Funktion F gibt so daß

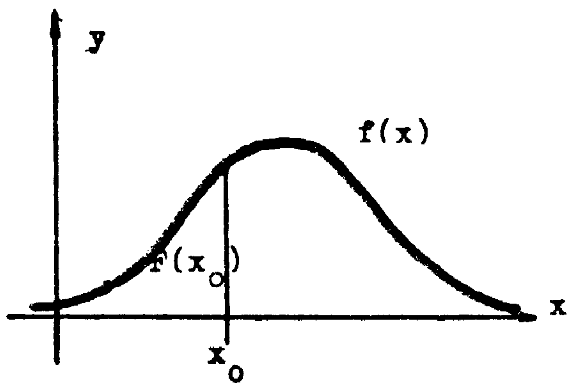
$$F(x) = \int_{-\infty}^{x} f(t)dt \quad \text{für alle} \quad x \in R.$$

Diese Definition vermeidet den relativ komplizierten Begriff der Differenzierbarkeit, man kann daraus aber nicht so leicht die (heuristische) Bedeutung der Funktion $f(x)$ erschließen. Da nach dem Hauptsatz der Differential- und Integralrechnung die Funktion dann stetig ist, heißt jede Zufallsvariable X, deren Verteilungsfunktion eine Dichte hat, auch stetige Zufallsvariable. Es muß jedoch nicht jede kontinuierliche Zufallsvariable auch stetig sein.

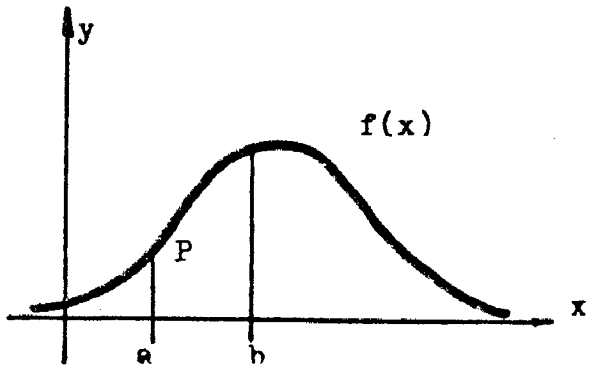


Da das bestimmte Integral als Fläche interpretiert werden kann

Darstellung von $F(x)$ als Flächeninhalt:



Darstellung von $P(a < X \leq b)$ als Flächeninhalt:



erhält man einen anschaulichen Zusammenhang zwischen Dichtefunktion und Verteilungsfunktion: $F(x)$ ist der Flächeninhalt des Normalbereichs zwischen $X = -\infty$ und $X = x$. Insbesondere ist für $b > a$

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$$

Für die Anwendungen ist praktisch nur diese Gleichung wichtig!

F ist also eine Stammfunktion von f (vgl. den "Hauptsatz" der Differential- und Integralrechnung).

Bemerkung 1: Bei kontinuierlichen Zufallsvariablen, die eine Dichtefunktion haben, dient diese Funktion z.B. zur Prognose der zu erwartenden relativen Häufigkeiten: $h_n(X = a)$ ist proportional zu $f(a)$. Wenn die Dichtefunktion wie in unseren Skizzen einen Gipfel hat, so ist offenbar der x-Wert, wo das Maximum aufgenommen wird, derjenige Wert, der mit größter Wahrscheinlichkeit (d.h. auf lange Sicht am häufigsten) angenommen wird.

Bei <u>diskreten</u> Zufallsvariablen zieht man für eine Prognose ("Schätzung") der relativen Häufigkeiten den Wert $P(X=a) \approx h_n(X=a)$ heran. Bei kontinuierlichen Zufallsvariablen versagt dies aber zumeist, weil ja im allgemeinen für jedes $a: P(X=a)=0$ gilt, wobei aber zu bedenken ist, daß schon die Fragestellung nicht angemessen ist, denn:

<u>Bemerkung 2:</u> Im übrigen entspricht die Verteilungsfunktion F einer Zufallsvariablen X weitgehend den in der Praxis auftretenden Problemen. Zur Diskussion stehen ja z.B. nicht die Wahrscheinlichkeit einer bestimmten kritischen Belastung x_o, sondern die Frage nach der Wahrscheinlichkeit des Unterschreitens bestimmter Höchstgrenzen, also $F(x_o)$. Auch bei anderen Anwendungen der Wahrscheinlichkeitsrechnung ist nicht ein ganz bestimmter Normwert x_o von Interesse, sondern die Wahrscheinlichkeit, daß bestimmte Werte innerhalb vorgegebener Grenzen liegen : $P(a < x \leq b) = F(b) - F(a)$.

Um solche Wahrscheinlichkeiten zu berechnen, muß man also F bzw f kennen. Die Natur macht uns nun die "Freude", daß bei den "naturgegebenen" Beispielen praktisch nur einige ganz wenige und typische Verteilungsfunktionen auftreten (Binomialverteilung, hypergeometrische Verteilung, Normalverteilung, Poissonverteilung,...). Alle diese Funktionen sind hinreichend genau <u>tabelliert,</u> so daß man tatsächlich kaum langwierige Summen bzw. Integrale berechnen muß. Um die Tabellen allerdings richtig benützen zu können, muß man den oben geschilderten "Hintergrund" kennen.

Im folgenden zeigen wir anhand von Beispielen, wie Wahrscheinlichkeiten aus der vorgegebenen Dichtefunktion mit Hilfe von Flächeninhalten berechnet werden können.

Beispiel 7: Eine Großbäckerei geht erfahrungsgemäß davon aus, daß an Samstagen ihr Brotabsatz X (in 1000 kg) die Dichtefunktion f(x) besitzt:

$$f(x) = \begin{cases} 0{,}04x & \text{für } 0 \leq x \leq 5 \\ 0{,}4 - 0{,}04x & \text{für } 5 < x \leq 10 \\ 0 & \text{sonst} \end{cases}$$

Diese Dichtefunktion entspricht den relativen Häufigkeiten.
Wie groß ist die Wahrscheinlichkeit, daß

(a) mehr als 5000 kg Brot verkauft werden;

(b) der Brotabsatz zwischen 4000 kg und 6000 kg Brot liegt?

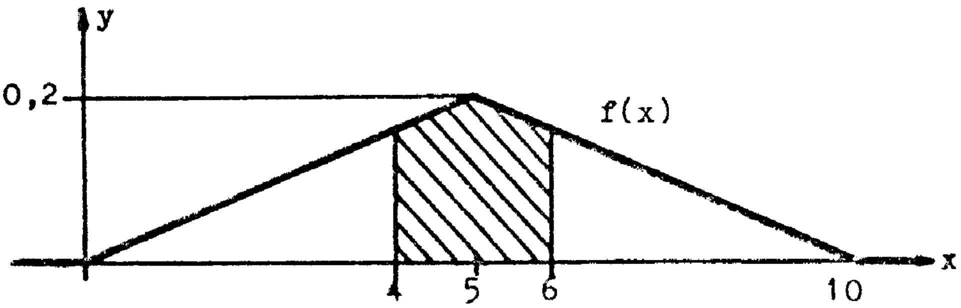
Lösung

Vorerst sollte man untersuchen, ob es sich dabei überhaupt um eine Dichtefunktion handelt.
Was muß denn dann gelten:

1) $f(x) \geq 0$

2) $\int_{-\infty}^{+\infty} f(x)dx = 1$

Am Graphen der Dichtefunktion, sieht man, daß 1) erfüllt ist und 2) erkennt man aus $A = \frac{10 \cdot 0{,}2}{2} = 1$



Nach unseren Überlegungen kann man die eingeschlossene Fläche als Wahrscheinlichkeit interpretieren und erhält so:

(a) $P(X > 5) = 0{,}5$, da ja ebenso $P(X \leq 5) = 0{,}5$ ist (Symmetrie!).

(b) $P(4 < X \leq 6) = 2 \cdot \frac{0{,}2 + 0{,}16}{2} \cdot 1 = 0{,}36$

4. Näherungswerte für große Fakultäten

Beispiel 1 (Abänderung): Wir interessieren uns für die Wahrscheinlichkeit, daß bei zufälliger Auswahl genau 17 von den 183 Abgeordneten weiblich sind.

Lösung: Binomialverteilung mit $p = 0{,}5$, $n = 183$ und $x = 17$.

$$P(X=17) = \binom{183}{17} \cdot 0{,}5^{17} \cdot 0{,}5^{183-17} = \binom{183}{17} \cdot 0{,}5^{183}$$

Für $\binom{183}{17}$ wird man aber keine Tabelle finden. Will man den Binomialkoeffizienten $\binom{183}{17}$ berechnen, so treten im Zähler 16 Multiplikationen auf. Auch die Umwandlung $\binom{183}{17} = \frac{183!}{17!\,166!}$ hilft nicht weiter, da n! für $n \geq 69$ mit dem Taschenrechner (Fakultätstaste) nicht mehr berechnet werden kann. Es ist daher sinnvoll, eine Näherungsformel für n ! herzuleiten, die man benützt, wenn n groß ist.

Bei der Herleitung dieser Formel handelt es sich um ein schönes Beispiel "experimenteller Mathematik" (Stichwort: Heuristik!), wie es sich für den MU bei Gelegenheit bestens eignet. (Immer wieder muß betont werden, daß Mathematik nicht im "Gebrauch von vorgefertigter Mathematik" und allenfalls deren Anwendung besteht!). Die Schüler müssen allerdings erfahren, daß solcherart nur Vermutungen (oder Näherungsresultate) entstehen können, die einer Absicherung bedürfen. (Beweisbedürfnis!). Hiezu entwickeln die Mathematiker dann Theorien, die in Beweisen oder Widerlegungen münden können. Für den MU eignet sich diesbezüglich oft auch eine bloße "Mitteilung", ein Kurzbericht aus dem Hintergrundwissen des Lehrers. (Siehe auch HANISCH, 1984b)
n ! bezeichne wie üblich das Produkt $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$

Um zu einer Näherung zu kommen, kann man etwa versuchen, die n verschiedenen Faktoren durch dieselbe Anzahl gleicher Faktoren zu ersetzen. Wählen wir versuchsweise etwa als den gleichen Faktor den Wert $\frac{n}{2}$, so erhalten wir folgende Näherung, wobei wir das "Korrekturglied" A(n) nicht kennen: $n! \approx \left(\frac{n}{2}\right)^n \cdot A(n)$

Um die Zahlen kleiner zu halten, logarithmieren wir:

$$\ln n! = \underbrace{n\ \ln - n \ln 2 + \ln A(n)}_{} = n \ln n + B(n)$$

B(n)

Untersuchen wir diese Näherung:

n	ln n ! − n ln n = B (n)
10	− 7,921
20	−17,579
30	−27,377
40	−37,235
50	−47,123
60	−57,033
70	−66,957

Offensichtlich ist $B(n) \sim 3-n$; somit ist $\ln n! \approx n \cdot \ln n \cdot (3-n)$ bzw. entlogarithmiert:

$$n! \approx e^{n.\ln n} . e^{3-n} = n^{n} . e^{3-n}$$

Durch Einführung eines neuen Korrekturgliedes C(n) bzw. D (n) kann man diese "Näherungsgleichung" als Gleichung schreiben:

$$n! = n^{n} \cdot e^{-n} \cdot \underbrace{e^{3} \cdot C(n)}_{D(n)}$$

Untersuchen wir $D(n) = \dfrac{n!}{n^{n} \cdot e^{-n}}$

n	$D(n) = \dfrac{n!}{n^{n} \cdot e^{-n}}$	Division durch 7,993	
10	7,993	1	$\approx \sqrt{1}$
20	1,257	1,41	$\approx \sqrt{2}$
30	13,771	1,72	$\approx \sqrt{3}$
40	15,886	1,99	$\approx \sqrt{4}$
50	17,754	2,22	$\approx \sqrt{5}$
60	19,433	2,43	$\approx \sqrt{6}$
70	20,997	2,63	$\approx \sqrt{7}$

Offenbar ist $D(n) \approx D(10) \cdot \frac{n}{20}$, also $D(n) \approx K \cdot \sqrt{n}$, wobei

$K = \frac{D(n)}{\sqrt{n}}$ eine Konstante oder - wenigstens für nicht allzu große n - angenähert konstant ist. Dies "prüfen" wir, indem wir diesen Wert $\frac{D(n)}{\sqrt{n}}$ (wir nennen ihn E(n)) für verschiedene n berechnen:

n	E(n)
10	2,53
20	2,52
30	2,52
40	2,51
50	2,51
60	2,51
70	2,51

Offenbar "pendelt" sich E(n) bei einem Wert ein, der - zumindest für Zahlen zwischen 40 und 70 mit 2,51 angenommen werden darf. Somit haben wir - jedenfalls für solche n - eine Näherungsformel für n ! gefunden, nämlich:

$$n! \approx 2{,}51 \sqrt{n} \cdot n^n \cdot e^{-n} = \sqrt{6{,}30 \cdot n} \cdot n^n \cdot e^{-n}$$

die leicht mit dem TR oder logarithmisch berechnet werden kann. Exakte und tieferliegende Überlegungen zeigen, daß diese Formel tatsächlich für beliebig große n verwendet werden kann und sogar umso besser ist, je größer n ist. Statt 6,30 ergibt sich bei diesen Überlegungen der Wert $2\pi = 6{,}283...$

Es gilt die sogenannte STIRLINGsche Formel:

$$n\,! \approx \sqrt{2\pi\, n}\;\; n^n \cdot \; e^{-n}$$

Bemerkung: Diese Formel ist in Wahrheit eine Grenzwertansage. Man kann nämlich beweisen, daß $\lim\limits_{n \to \infty} \frac{n}{\sqrt{2\pi n}\; n^n e^{-n}} = 1$ gilt.

(Stichwort: Zähler und Nenner sind asymptotisch gleich)
Um zu überprüfen, wie gut die Näherung ist, verwenden wir den von SIEBER in den Mathematischen Tafeln angegebenen Wert:

$100! = 9{,}33262 \cdot 10^{157}$

$100! = \sqrt{2\pi\, 100} \cdot 100^{100}\, e^{-100} = 9{,}325 \cdot 10^{157}$

Wir sehen, der Fehler ist weniger als 1 %.

Bemerkung: 100! läßt sich natürlich nicht so ohne weiteres mit dem TR berechnen, wenn man aber dekadisch logarithmiert:

$\lg 100! \approx \lg \sqrt{2\pi 100} + 200 - 100 \lg e =$
$= 157{,}96964 = 157 + 0{,}96964$ geht es.

Sicherlich wird man sich fragen, ob die gebrachte Herleitung auch wirklich eine "Herleitung" ist und nicht eine "Herumprobiererei". Nun sie ist typisch für die Forschungsarbeit eines Mathematikers - beides: Vgl. das eingangs Gesagte. Freudenthal meinte einmal: "Einem Kind ein Geheimnis zu verraten, das es selber entdecken kann, ist schlechte Didaktik, es ist ein Verbrechen". Selbstverständlich muß man bei einem solchen experimentellen Unterricht auch riskieren, daß manche Tätigkeiten und Einfälle <u>nicht</u> zum Ziel führen, auch darin erblicke ich einen gewissen erzieherischen Wert. Jedenfalls liegt in der Erstellung von Tabellen und deren experimentierender Auswertung ein wesentliches Lehrziel, insbesondere dann, wenn es sich - wie hier - um "gezieltes" und "geplantes" Probieren handelt.

Nun können wir darangehen, unser Problem nicht nur für genau $x = 17$ Frauen, sondern wie aus der ursprünglichen Fragestellung hervorgeht für $x = 0, 1, \ldots 17$ lösen. Am besten in einer Gruppenarbeit

Lösung von Beispiel 1:

$P(X \le 17) = \sum_{x=0}^{17} \binom{183}{x} \cdot 0{,}5^{183} = 0{,}5^{183} \sum_{x=0}^{17} \binom{183}{x} =$

$= 8{,}157 \cdot 10^{-56} \cdot 4{,}207 \cdot 10^{23} = 3{,}431 \cdot 10^{-32} \approx 0$

x	$\binom{183}{x}$
0	1
1	$1{,}83 \cdot 10^{2}$
2	$1{,}6653 \cdot 10^{4}$
3	$1{,}0047 \cdot 10^{6}$
4	$4{,}5213 \cdot 10^{7}$
5	$1{,}6186 \cdot 10^{9}$
6	$4{,}8019 \cdot 10^{10}$
7	$1{,}2142 \cdot 10^{12}$
8	$2{,}6712 \cdot 10^{13}$
9	$5{,}1941 \cdot 10^{14}$
10	$9{,}0377 \cdot 10^{15}$
11	$1{,}4214 \cdot 10^{17}$
12	$2{,}0373 \cdot 10^{18}$
13	$2{,}6798 \cdot 10^{19}$
14	$3{,}2541 \cdot 10^{20}$
15	$3{,}6663 \cdot 10^{21}$
16	$3{,}8496 \cdot 10^{22}$
17	$3{,}7817 \cdot 10^{23}$
	$4{,}2068 \cdot 10^{23}$

Einfacher ist allerdings folgende Überlegung:

$P(X=17)=\binom{183}{17} \cdot 0{,}5^{183}=3{,}7817 \cdot 10^{23} \cdot 8{,}157 \cdot 10^{-56} = 3{,}08 \cdot 10^{-32}$

Da die anderen Wahrscheinlichkeiten alle kleiner sind, muß gelten

$P(X \le 17) \le 17 \cdot 3{,}08 \cdot 10^{-32} = 5{,}24 \cdot 10^{-31} \approx 0$

Die Wahrscheinlichkeit ist also praktisch Daraus folgt, daß unsere Annahme "Eine Frau hat die gleiche Chance, in den Nationalrat entsendet zu werden" falsch ist. Es liegt daher eine Diskriminierung der Frauen vor.

5. Herleitung der Normalverteilung

Die 17 Binomialkoeffizienten lassen sich in einer Gruppenarbeit noch leicht berechnen, aber stellen Sie sich vor, auf Grund dieser Arbeit werden nach den nächsten Wahlen fünfmal so viele Frauen als Ageordnete in den Nationalrat (also 5.17=85) entsendet (Beispiel 1a).
Da.wäre die Berechnung der Binomialkoeffizienten schon sehr aufwendig und um die auftretende Rechenarbeit zu erleichtern, ist es günstig, eine Näherungsformel zu suchen. Dabei gehen wir von folgender Idee aus:

Binomialverteilungen können durch Histogramme dargestellt werden, wobei der Flächeninhalt gleich der Wahrscheinlichkeit ist. Wenn die Abszissenabschnitte nur genügend klein gewählt werden, müßte es möglich werden, die Treppenfunktion durch eine Kurve anzunähern, wodurch das praktische Rechnen erleichtert würde.

Um Binomialverteilungen miteinander vergleichen zu können, führt man eine Standardisierung durch, und benützt dazu den Mittelwert μ und die Standardabweichung σ zur folgenden Substitution von x durch z :

$$z = \frac{x - \mu}{\sigma}$$

Der Erfolg dieser Kombination einer Parallelverschiebung und einer Streckung ist eine Verteilung mit dem Mittelwert 0 und der Streuung 1.

Beweis: $E(Z) = E\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma} E(x) - \frac{\mu}{\sigma} = 0$

$$\sigma^2(Z) = \sigma^2\left(\frac{x - \mu}{\sigma(x)}\right) = \frac{1}{\sigma^2(x)} \sigma^2(x-\mu) = \frac{1}{\sigma^2(x)} \cdot \sigma^2(x) = 1$$

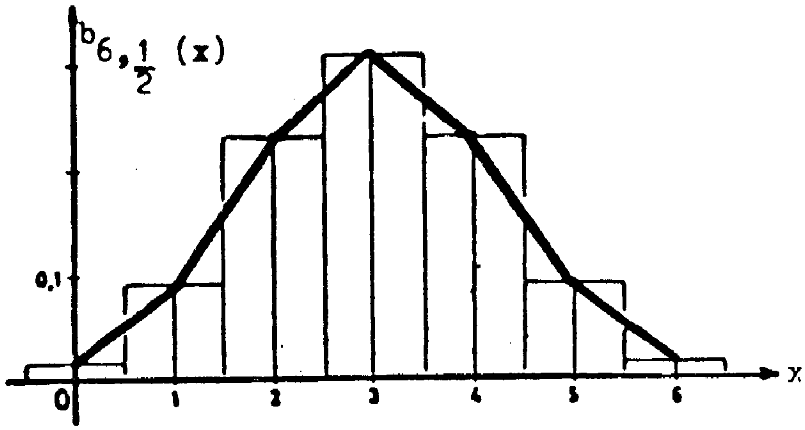
Durch diese Transformation (Standardisierung) geschieht folgendes:

1) Der Nullpunkt fällt jetzt mit dem Mittelwert μ zusammen.
2) Die Breite der Treppenstufen wird verändert. Die neue Breite ist das $\frac{1}{\sigma}$ -fache der alten.

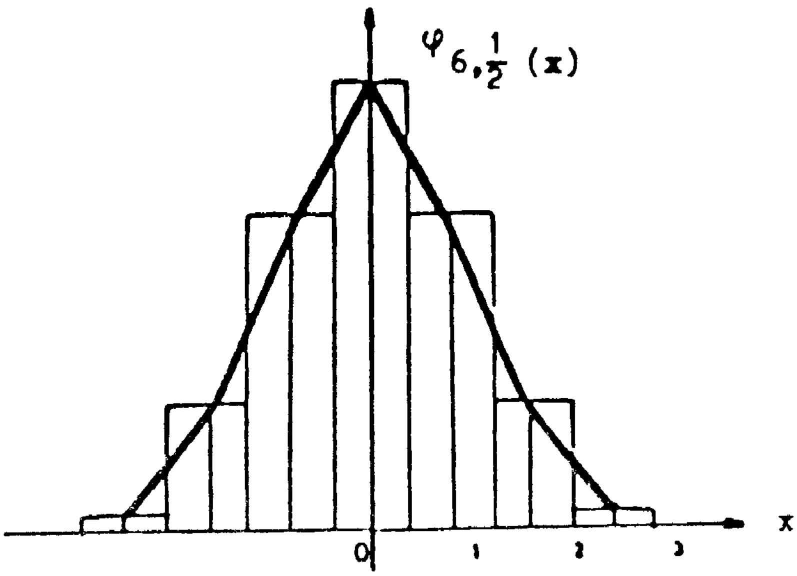
Damit aber die Summe der Flächen aller Treppenstufen wieder 1 ist (Fläche ist Maß für die Wahrscheinlichkeit), müssen die Ordinaten zum Ausgleich mit σ multipliziert werden:

$$\varphi(z) = f(x) \cdot \sigma$$

Wählen wir zur Demonstration eine Binomialverteilung mit $n = 6$ und $p = \frac{1}{2}$, so ist $b_{6,\frac{1}{2}}(x)$:



Da $\mu = n.p = 3$ und $\sigma = \sqrt{n.p(1-p)} = \sqrt{1,5} = 1,22$ ist, ergibt sich, $z = \frac{x-3}{1,22}$ und $\varphi_{6,\frac{1}{2}}(z) = 1,22 \cdot b_{6,\frac{1}{2}}(x)$:



Allgemein gilt für beliebiges n und $p = \frac{1}{2}$:

$\mu = n.p = \frac{n}{2}$, $\sigma^2 = np(1-p) = \frac{n}{4}$

und $f(x) = \binom{n}{x} \cdot (\frac{1}{2})^x \cdot (\frac{1}{2})^{n-x} = \binom{n}{x} \cdot \frac{1}{2^n} = \frac{n!}{x!\,(n-x)!} \cdot \frac{1}{2^n}$

Untersuchen wir nun, was mit dieser Binomialverteilung geschieht, wenn n sehr groß wird. Dazu verwenden wir die STIRLING-Formel für n!

Es ergibt sich $f(x) \approx \dfrac{n^n \cdot \sqrt{n}}{\sqrt{2\pi} \cdot \sqrt{x(n-x)} \cdot x^x (n-x)^{n-x}} \cdot \dfrac{1}{2^n}$

Diese Transformation $x \to z$ hat nun die Gestalt

$$z = \frac{x-\mu}{\sigma} = \frac{2}{\sqrt{n}}\,(x - \frac{n}{2}) = \frac{2\,x-n}{\sqrt{n}}$$

woraus folgt

$x = \frac{1}{2}\ (n+z \cdot \sqrt{n})$ und $n-x = \frac{1}{2}\,(n-z \cdot \sqrt{n})$

Somit ist

$\varphi(z) = \sigma \, . \, f(x) = \frac{\sqrt{n}}{2} \cdot f(x) =$

$$= \frac{\sqrt{n} \cdot n^n \cdot \sqrt{n} \qquad 2^{\frac{n+z\sqrt{n}}{2}} \cdot 2^{\frac{n-z\sqrt{n}}{2}}}{2 \cdot \sqrt{2\pi} \cdot \sqrt{\frac{1}{4}\,(n^2-z^2n)} \cdot (n+z\sqrt{n})^{\frac{n+z\sqrt{n}}{2}} \cdot (n-z\sqrt{n})^{\frac{n-z\sqrt{n}}{2}}} \cdot \frac{1}{2^n} =$$

Durch Zusammenfassen und Kürzen erhält man:

$$\varphi(z) = \sqrt{\frac{n^2}{n^2-z^2n}} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{n^n}{(n^2-z^2n)^{\frac{n}{2}}} \cdot \left(\frac{n-z\sqrt{n}}{n+z\sqrt{n}}\right)^{\frac{z\sqrt{n}}{2}} =$$

$$= \sqrt{\frac{1}{1-\frac{z^2}{n}}} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{(1-\frac{z^2}{n})^{\frac{n}{2}}} \cdot \frac{\left[(1-\frac{z}{\sqrt{n}})^{\sqrt{n}}\right]^{\frac{z}{2}}}{\left[(1+\frac{z}{\sqrt{n}})^{\sqrt{n}}\right]^{\frac{z}{2}}}$$

Nun ist aber

$\lim\limits_{n\to\infty} (1\pm\frac{a}{n}) = e^{\pm a}$, daher geht $\varphi(z)$ für großes n über in

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{e^{-\frac{z^2}{2}}} \cdot \frac{e^{-z \cdot \frac{z}{2}}}{e^{z \cdot \frac{z}{2}}} = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}$$

Ist nun eine zufällige Größe so verteilt, daß ihre Wahrscheinlichkeitsdichte

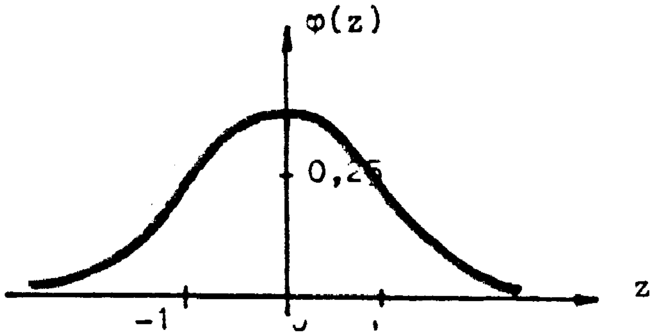
$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

ist, dann heißt sie normalverteilt mit dem Mittelwert 0 und der Varianz 1.

<u>Bemerkung:</u> Dieser Teil der Herleitung ist relativ aufwendig. Die Grenzwerte aber lassen sich in Arbeitsgruppen berechnen, und für den anderen Teil würde ich im MU die Seite kopieren und durch die Schüler verifizieren lassen. Eventuell baue man ein oder zwei Fehler ein (z.B. statt z^2n etwa zn) und fordere die Schüler auf, diese zu suchen.

Hat man nun die Funktionsgleichung der Dichtefunktion, so gibt eine Kurvendiskussion (endlich eine sinnvolle Anwendung einer solchen) Näheres über die Gestalt ihres Graphen:

1) $z \to -z \quad f(z) = f(-z)$	Die Kurve ist symmetrisch bezüglich der Ordinate.
2) $f'(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}} \cdot (-z)$	Ihr Extrempunkt ist bei z=0
3) $f''(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}} (-1+z^2)$	Sie besitzt zwei Wendepunkte bei $z = \pm 1$
4) $\lim\limits_{z\to\infty} f(z) = 0$	Die z-Achse ist Asymptote



Diese Funktion heißt nach ihrem Entdecker und ihrer Gestalt auch Gauß'sche Glockenkurve.

5. Aufstellen der Tabelle für $\Phi(x)$

Leider kann man das Integral $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt = \Phi(z)$ nur näherungsweise berechnen, daher ist es günstig, eine Tabelle dafür aufzustellen. Auch hier bietet sich wieder eine Gruppenarbeit an.

Aus Symmetriegründen ist $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{t^2}{2}} dt = 0{,}5$ und wir können uns auf Weite für $z > 0$ beschränken. Wollen wir beispielsweise eine Tabelle in Schritten von 0,1 erstellen, so können wir, wenn wir $\Phi(z)$ kennen, $\Phi(z+0{,}1)$ erhalten durch:

$$\Phi(z+0{,}1) = \Phi(z) + \frac{1}{\sqrt{2\pi}} \int_{z}^{z+0{,}1} e^{-\frac{t^2}{2}} dt$$

wobei das letztere Integral leicht mittels der KEPLERschen Faßregel ausgewertet werden kann:

$$\int_a^b f(x)dx = \frac{b-a}{6} \left[f(a) + 4\cdot f\left(\frac{a+b}{2}\right) + f(b)\right]$$

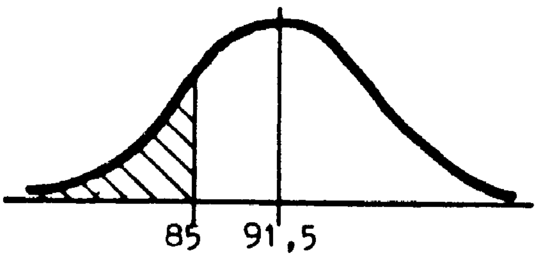
Man verifiziert damit leicht die in CESNEK-FLODERER (1980) angegebenen Werte.

Lösung von Beispiel 1a: Nun wollen wir (endlich) unser Beispiel lösen:

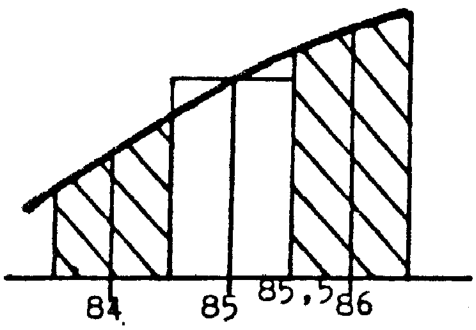
$n = 183 \quad p = 0{,}5 \quad x = 85$

$\mu = n.p = 91{,}5, \sigma = \sqrt{183 \cdot 0{,}5 \cdot 0{,}5} = 6{,}76$

Stellen wir das in einer Zeichnung dar:



Zu beachten ist aber noch, daß die Binomialverteilung eine diskrete Verteilung ist, sie aber durch die kontinuierliche Normalverteilung angenähert wird.
Betrachten wir dies genauer bei x = 85 (Funktionenmikroskop):



Wir sehen: Würde als Integrationsgrenze bei der Normalverteilung x=85 gewählt werden, so wäre die Näherung sicher nicht so gut, als wenn x = 85,5 gewählt wird. Wir haben daher bis 85,5 zu integrieren.

Die untere Grenze kann hier vernachlässigt werden, da der dabei begangene Fehler praktisch Null ist.
Die Standardisierung liefert:

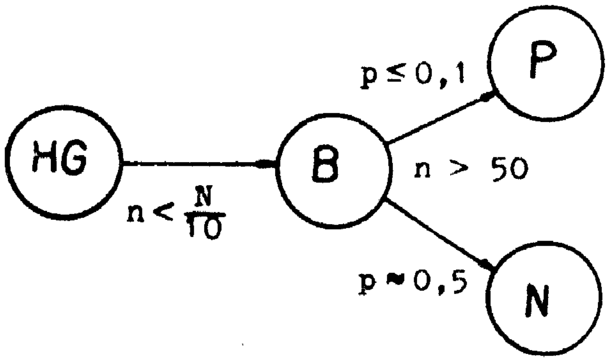
$$z = \frac{x - \mu}{\sigma} = \frac{85,5 - 91,5}{6,76} = -0,8876$$

Wir brauchen also Φ (-0,89). In unserer Tabelle finden wir aber nur $\Phi(0,89)$ = 0,81. Aus Symmetriegründen ist $\Phi(0,89)$ =
= 1 - $\Phi(0,89)$ = <u>0,19</u>.

Die Wahrscheinlichkeit, daß bei gleichen Chancen von 183 Abgeordneten 85 oder weniger Frauen sind, ist 19%. Im allgemeinen nimmt man bei einer so hohen Wahrscheinlichkeit nicht mehr an, daß das Ergebnis unwahrscheinlich ist, es gäbe also bei einer so hohen Zahl von weiblichen Abgeordneten keinen Anlaß, von einer Diskriminierung zu sprechen.

Bemerkung: Didaktisch günstig wäre es, in den saueren Apfel zu beißen und in Gruppenarbeit das Beispiel auch exakt mit der Binomialverteilung durchzurechnen, um einen Vergleichswert zu haben und zu sehen, wie gut die Näherung ist.
Untersucht man schließlich noch Binomialverteilungen, bei denen $p \neq 0,5$ ist, so zeigt sich, daß die angegebene Näherung auch bei dieser brauchbar ist, soferne p nicht allzu klein und n genügend groß ist.

Allgemein gilt folgende Faustregel für die Anwendbarkeit und den Zusammenhang zwischen den wichtigsten Verteilungen:



Diese wird man natürlich um soweit bringen, wie man die Hypergeometrische und Poissonverteilung besprochen hat. Die Verifikation wäre ein lohnendes Beispiel für den EDV-Unterricht.
Ansonsten kann man auch an einem Beispiel mit sehr kleinem p zeigen, daß die Approximation durch die Normalverteilung ganz brauchbar ist.

Beispiel 8: Eine Lieferung wird abgelehnt, wenn in einer Stichprobe von 500 Stück mindestens 5 defekte Stücke enthalten sind. Sind tatsächlich 0,5% der Stücke defekt, wie groß ist dann die Wahrscheinlichkeit, daß die Lieferung angelehnt wird?
Die Wahrscheinlichkeit ist auf der Basis der

a) Binomialverteilung
b) Poissonverteilung
c) Normalverteilung

zu berechnen.

<u>Lösung</u>:

a) $n = 500$ $\qquad p = 0{,}005$

Gesucht ist $P(x \geq 5)$. Es ist besser mit der Gegenwahrscheinlichkeit zu rechnen.

$$P(x \geq 5) = 1 - P(x \leq 4)$$

$$P(X = x) = \binom{500}{x} \cdot 0{,}005^{x} \cdot 0{,}995^{500-x}$$

x	$f_B(x)$	$f_p(x)$
1	0,0816	0,0821
1	0,2050	0,2052
2	0,2570	0,2565
3	0,2144	0,2138
4	0,1338	0,1336
Σ	0,8918	0,8912

$P(X \geq 5) = \underline{\underline{0{,}108}}$

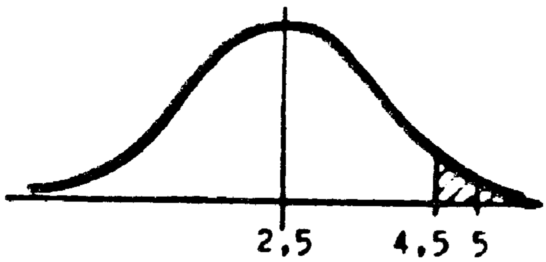
b) $\mu = n.p = 500 \cdot 0{,}005 = 2{,}5$

$$P(X \geq 5) = 1 - P\,(X \leq 4)$$

$$P(X = x) = f_p(x) = \frac{2{,}5^{x}}{x!} \quad e^{-2{,}5}$$

$P(X \geq 5) = \underline{\underline{0{,}109}}$

c) $\sigma^2 = n.p.(1-p) = 2{,}5 \cdot 0{,}995 = 2{,}4875$



$$z = \frac{x-\mu}{\sigma} = \frac{4{,}5-2{,}5}{1{,}577} = 1{,}268$$

$$\Phi(1{,}268) = 0{,}898$$

$P(X > 5) = \underline{\underline{0{,}102}}$

Wir sehen an diesem Beispiel, daß die Näherung durch die Normalverteilung doch ganz brauchbar ist, obwohl p weit von 0,5 abweicht. Trotzdem wäre es ein Kunstfehler, nicht die Poissonverteilung zu verwenden (soferne diese im Unterricht besprochen wurde).

Bei den meisten Anwendungen der Normalverteilung wird allerdings nic eine Binomialverteilung angenähert, sondern man weiß aus Erfahrung, daß die Zufallsvariable normalverteilt ist.

Beispiel 6 (Fortsetzung): Auf einer Drehbank werden Werkstücke abgedreht, deren Durchmesser 200 mm betragen soll und dessen tatsächlicher Wert naturgemäß schwankt. Er ist daher eine zufällige Größe und - wie man weiß - normalverteilt mit dem Mittelwert 200 mm (hängt von der Maschineneinstellung ab) und der Standardabweichung 0,02 mm (hängt von der Qualität der Maschine ab).

Wieviel Prozent Ausschuß ist zu erwarten, wenn
a) der Durchmesser der Werkstücke mindestens 199,95 mm stark sein soll
b) der Durchmesser zwischen 199,97 und 200,03 mm liegen muß.

Lösung:

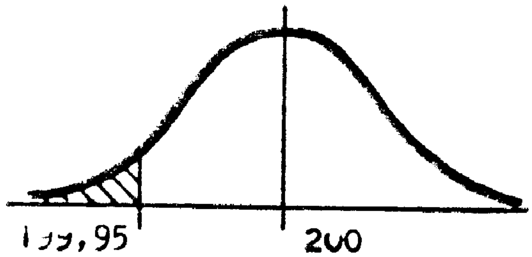
a) X...Durchmesser in mm

$f(X)$ ...N(200; 0,02)

$$P(X \le 199{,}95) = \Phi\left(\frac{199{,}95 - 200{,}0}{0{,}02}\right) =$$

$$= \Phi\left(-\frac{0{,}05}{0{,}02}\right) = \Phi(-2{,}5) = 0{,}0062 =$$

$$= \underline{\underline{0{,}62\ \%}}$$

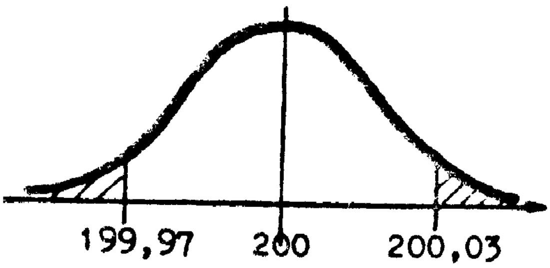


b) $$1-P(199{,}97 < X \le 200{,}03) = 1 - \left[\Phi\left(\frac{200{,}03 - 200}{0{,}02}\right) - \Phi\left(\frac{199{,}97-200}{0{,}02}\right)\right] =$$

$$= 1 - \left[\Phi\left(\frac{0{,}03}{0{,}02}\right) - \Phi\left(-\frac{0{,}03}{0{,}02}\right)\right] =$$

$$= 1 - [\Phi(1{,}5) - \Phi(-1{,}5)] =$$

$$= 1 - 0{,}8664 = 0{,}1336 = \underline{\underline{13{,}36\%}}$$



6. Zusammenfassung

Hat man vor, die Normalverteilung im Unterricht zu besprechen, so ist es m.E. didaktisch günstig schon sehr bald den Begriff der "relativen Häufigkeitsdichte" einzuführen, der sich beim

Zeichnen von Histogrammen zwanglos anbietet. Der weitere Weg führt dann über die kontinuierlichen Verteilungen und den Begriff der Zufallsvariablen zur Verteilungs- und zur Dichtefunktion, die eine lohnende Anwendung für die Integralrechnung bieten. Um den Binomialkoeffizienten $\binom{n}{k}$ auch für große n und k berechnen zu können, wird anschließend experimentell die STIRLING'sche Formel entdeckt. Diese wird dann dazu verwendet, um für den Spezialfall p = o,5 aus der Binomialverteilung die Normalverteilung (in teilweiser Gruppenarbeit) herzuleiten. Anschließend wird wieder in Gruppenarbeit die Tabelle für die Verteilungsfunktion der Normalverteilung hergeleitet und empirisch gezeigt, daß die für p = 0,5 hergeleitete Näherung auch für andere p brauchbar ist. Daran könnte sich nun eine Diskussion des zentralen Grenzwertsatzes anschließen, um die Bedeutung der Normalverteilung aufzuzeigen.

Abschließend möchte ich noch den Herren Doz.Dr.Bürger, Mag.R. Müller und vor allem Prof.Dr.H.G.Reichel danken, die mich bei der Erstellung dieser Arbeit unterstützten.

7. Literatur

Hier ist nur die in dieser Arbeit zitierte Literatur aufgenommen.

CESNEK.E. u. FLODERER. M.: Tabellen. Zahlentafeln-Einheiten-Formeln. Wien, 1980.

HANISCH, G.: Sum Derivations. In: The Mathematics Teacher, vol. 77,Nr.9, Dec. 1984b.

——— : Mathematik und Moral. In: GROSSER, S. (Hrsg): Didaktik-Reihe der ÖMG, Heft 11, 1984a.

——— gem. mit REICHEL, H.G.: Didaktik der Stochastik (Arbeitstitel). Wien, 1986 (im Druck).