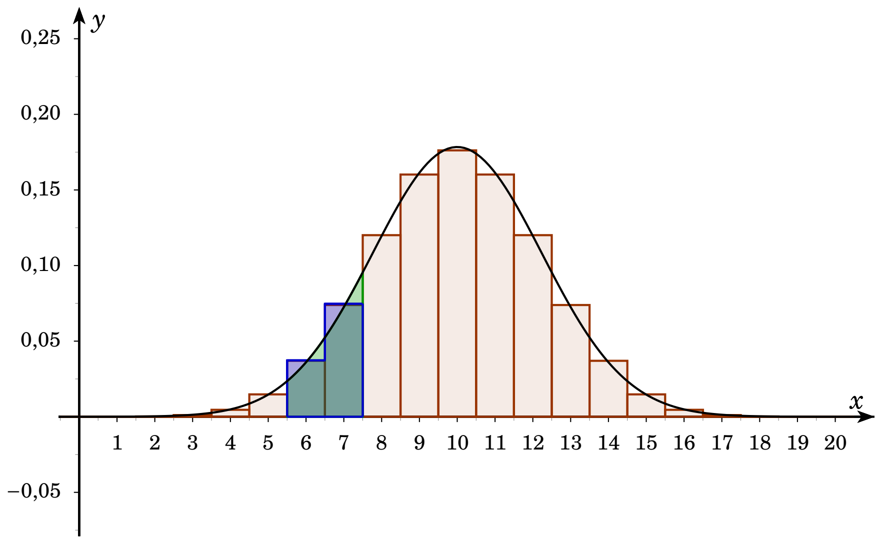
Freie Hansestadt Bremen Gymnasium LdW	Klasse Q1a *Datum:*	Mathematik 12.08.2014	Lehrer: Torsten Warncke

1. Binomial- und Normalverteilung



1.1. Binomialverteilung

Obige Graphik zeigt eine Binomialverteilung $B(20;0{,}5)$ und eine entsprechende Normalverteilung $N(10;\sqrt{5})$.
Hierbei ist das rötlich unterlegte Histogramm die Darstellung der Binomialverteilung. Der häufigste Wert (Modalwert, Modus) liegt bei 10, er ist gleichzeitig auch der Medianwert, Mittelwert und Erwartungswert $E(x) = n \cdot p = \mu = 10$. Symmetrisch nehmen die Säulen um μ ab, bis sie für $x = 0$ und $x = 20$ praktisch keine sichtbare Höhe mehr haben. Die Säulen für $x = 6$ und $x = 7$ sind außerdem blau unterlegt. Ihre Flächen summieren sich zu dem Wert

$$P(6 \leq x \leq 7) = P(x = 6) + P(x = 7) \approx 0{,}111, \tag{1}$$

diesem Flächenwert entspricht die Wahrscheinlichkeit, dass die diskret verteilte Zufallsvariable x die Merkmalsausprägungen 6 oder 7 annimmt, wenn der Zufallsversuch insgesamt $n = 20$ mal wiederholt wird und die Trefferwahrscheinlichkeit $p = 0{,}5 = 50\%$ beträgt.
Rechnerisch handelt es sich um

$$\mathrm{binomCDF}(20;0{,}5;7) - \mathrm{binomCDF}(20;0{,}5;5) = \sum_{k=6}^{7} \binom{n}{k} p^k (1-p)^{n-k} = \tag{2}$$

$$= \binom{20}{6} 0{,}5^6 (1-0{,}5)^{20-6} + \binom{20}{7} 0{,}5^7 (1-0{,}5)^{20-7} = 0{,}0369644 + 0{,}0739288 \approx 0{,}111, \tag{3}$$

mit binomCDF$(n;p;m) = \sum_{k=0}^{m} \binom{n}{k} p^k (1-p)^{n-k}$.

1.2. Normalverteilung

Analog verhält sich die Normalverteilung, deren Graph (siehe 1) durch eine schwarze Linie dargestellt ist, die sogenannte Gaußsche Glockenkurve. Ihre Parameter $\mu = 10$ und $\sigma = \sqrt{5} \approx 2{,}2361$ wurden aus der Binomialverteilung wie folgt entnommen:

$$\mu = n \cdot p = 20 \cdot 0{,}5 \qquad = 10 \tag{4}$$

$$\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{20 \cdot 0{,}5 \cdot 0{,}5} = \sqrt{5} \qquad \approx 2{,}2361 \tag{5}$$

Um eine ähnliche Wahrscheinlichkeitsaussage wie für die Binomialverteilung zu erhalten, muss nun integriert (statt simpel aufsummiert) werden.

$$P(5{,}5 \le x \le 7{,}5) = \int_{5{,}5}^{7{,}5} \varphi(x) dx = \Phi(7{,}5) - \Phi(5{,}5) \qquad \approx 0{,}1097. \tag{6}$$

D.h. die grün unterlegte Fläche, die der Graph von

$$\varphi(x) = \frac{e^{(x-10)^2/10}}{\sqrt{10\pi}} \tag{7}$$

mit der x-Achse zwischen $x = 5{,}5$ und $x = 7{,}5$ einschließt ist fast so groß wie die blau unterlegte Fläche bei der Binomialverteilung. Dieser grün unterlegten Fläche entspricht die Wahrscheinlichkeit, dass die stetig verteilte Zufallsvariable x Werte zwischen 5,5 und 7,5 annimmt, wenn sie normalverteilt mit Mittelwert $\mu = 10$ und Standardabweichung $\sigma = \sqrt{5} \approx 2{,}2361$ ist.

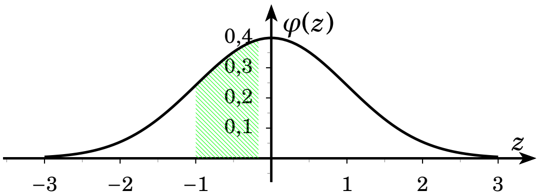
Allgemein wird die Normalverteilung beschrieben durch die Gauss-Funktion:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{8}$$

Hierbei ist μ der Mittel-, Modal-, Median- und Erwartungswert und σ die Standardabweichung.

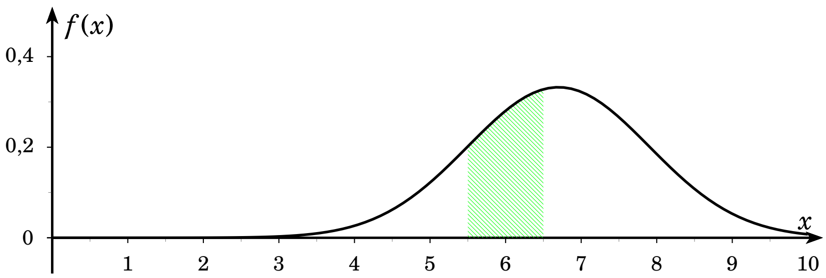
Nehmen wir einmal an, die Ergebnisse eines Biotests mit 10 Fragen seien normalverteilt (vgl. Schaum S. 132), die mittlere Punktzahl war 6,7 Punkte und die Standardabweichung 1,2.
Aufgabe: Bestimme den Prozentsatz der Prüflinge, der die Punktzahl 6 (6 von 10 Fragen richtig) hatte.
Lösung: Bei einer Normalverteilung muss alles stetig sein, d.h. die Punktzahl gehört zu den Prüflingen, die zwischen 5,5 und 6,5 Punkte erhielten. Standardisiert liegt dann $z = \frac{x-\mu}{\sigma}$ zwischen $z = \frac{5{,}5-6{,}7}{1{,}2} = -1$ und $z = \frac{6{,}5-6{,}7}{1{,}2} = -\frac{1}{6} \approx -0{,}17$, der gesuchte Prozentsatz

liegt mithilfe der Gauss-Tabelle bzw. mithilfe Integration (grünlich schraffierte Fläche) bei rund 27,5% (Schaums Musterlösung ist falsch).



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

$$\int_{-1}^{-0{,}17} \varphi(z)dz = \int_{5{,}5}^{6{,}5} f(x)dx \approx 0{,}275 = 27{,}5\%$$



Etwa 27,5% der Ergebnisse dieses Testes im Fach Biologie waren bei 6 Punkten (zwischen 5,5 und 6,5 von 10 möglichen Punkten).
Die zugehörigen Tabellenergebnisse könnten sein $\Phi(0{,}17) = 0{,}56749$ und $\Phi(1) = 0{,}84134$. Wegen der Achsensymmetrie der Gauss-Verteilung ist die Integralfunktion $\Phi(z)$ punktsymmetrisch zum Punkt $(0|0{,}5)$ und es gilt $\Phi(-z) = 1 - \Phi(z)$, also hier $\Phi(-0{,}17) = 1 - 0{,}56749 = 0{,}43251$ und $\Phi(-1) = 1 - 0{,}84134 = 0{,}15866$. Damit ergibt sich rechnerisch

$$\int_{-1}^{-0{,}17} \varphi(z)dz = \Phi(-0{,}17) - \Phi(-1) = 0{,}43251 - 0{,}15866 = 0{,}27385 = 27{,}385\% \approx 27{,}5\%.$$

Hier wurde „großzügig aufgerundet“, da das Integral als obere Grenze nicht $-0{,}17$ sondern eigentlich $-\frac{1}{6}$ (leider ohne Tabellenwert) hat und der Graph der Normalverteilung zwischen $-0{,}17$ und $-0{,}16$ noch kräftig ansteigt und „zusätzliche Fläche liefert“.

Abschließend die Graphen der (wieder schwarzen) Standardnormalverteilung $\varphi(z)$ mit deren (in blau) Stammfunktion $\Phi(z)$: