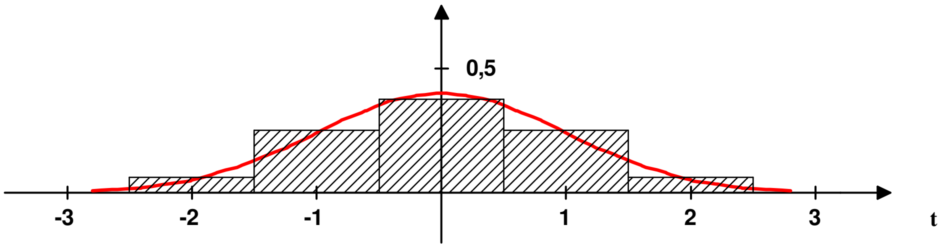
XIII. Die Normalverteilung

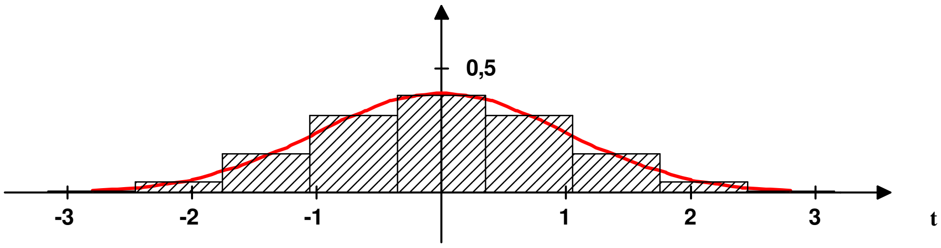
==

13.1 Der lokale Grenzwertsatz

Standardisierung von B(4; 0,5) :



Standardisierung von B(8; 0,5) :



Lokaler Grenzwertsatz :

Die Dichtefunktion φ_n (= Treppenfunktion, die ein Histogramm oben berandet) der standardisierten Binomialverteilung B(n; p) geht mit wachsendem n gegen die Grenzfunktion

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

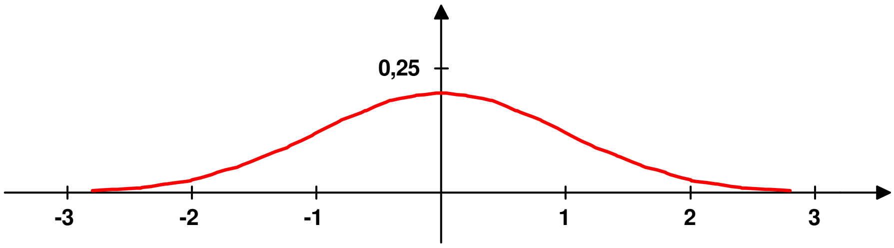
φ heißt deshalb Dichte der **Standardnormalverteilung**.

Eigenschaften von φ :

1. Achsensymmetrie $\boxed{\varphi(-t) = \varphi(t)}$

2. $\boxed{\lim\limits_{t \to \pm\infty} \varphi(t) = 0}$

3. $\boxed{\int_{-\infty}^{\infty} \varphi(t)dt = 1}$



Anwendung :

Für große n (npq > 9) gilt :

$$\boxed{\mathbf{B(n;\, p;\, k) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{1}{2}\frac{(k-np)^2}{npq}}}}\ \textbf{(lokale Näherung)}$$

Begründung :

Breite eines Histogrammrechtecks der standardisierten Binomialverteilung : $\Delta t = \frac{1}{\sigma} = \frac{1}{\sqrt{npq}}$

Höhe des zu k gehörenden Rechtecks : $h \approx \varphi(t) = \varphi(\frac{k-np}{\sqrt{npq}})$

$$\Rightarrow \ B(n;\, p;\, k) \approx h \cdot \Delta t = \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-\frac{1}{2}\left(\frac{k-np}{\sqrt{npq}}\right)^2}$$

Beispiel :

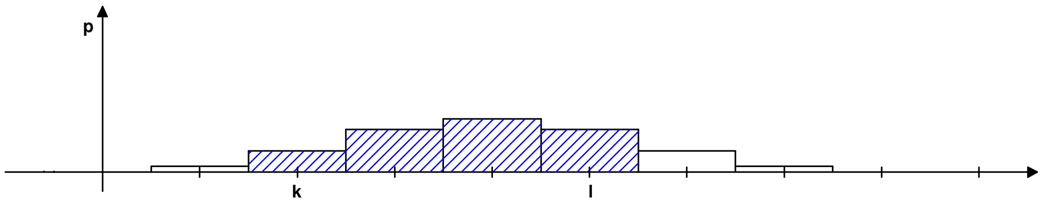
Es ist B(200; 0,5; 104) = 0,04805

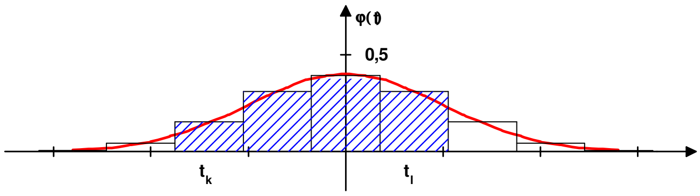
Die lokale Näherung ergibt : $B(200;\, 0{,}5;\, 104) \approx \frac{1}{\sqrt{2\pi\cdot 200\cdot 0{,}5\cdot 0{,}5}} e^{-\frac{1}{2}\frac{(104-100)^2}{200\cdot 0{,}5\cdot 0{,}5}} \approx 0{,}0410$

Bemerkung :

Eine Wertetabelle von φ findet man in der Stochastik-Tabelle (S.59).

13.2 Integraler Grenzwertsatz





Im Histogramm der Binomialverteilung B(n;p) ist die Flächensumme der Histogramm-rechtecke von k bis l mit $0 \le k \le l \le n$ gleich der Wahrscheinlichkeit $P(k \le X \le l)$ für die Trefferzahl X.

Ist X_S die Standardisierung von X und t_k und t_l die standardisierten Werte von k und l, dann

gilt

$P(t_k \le X_s \le t_l) \; = \; P(k \le X \le l)$.

Die Flächenberechnung lässt im Histogramm von X_S näherungsweise durch eine Integration durchführen.

$$P(k \le X \le l) \; = \sum_{k \le i \le l} B(n;\, p;\, i) \approx \int_{t_k}^{t_l} \varphi(t)dt \quad \int_{\frac{k-\mu-\frac{1}{2}}{\sigma}}^{\frac{l-\mu+\frac{1}{2}}{\sigma}} \varphi(t)dt$$

Setzt man

$$\Phi(x) := \int_{-\infty}^{x} \varphi(t)dt \; = \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt\,,$$

dann gilt

Satz :

Für die Binomialverteilung gilt folgende **Näherungsformel**

$$\boxed{\mathbf{P(k \le X \le l) \approx \Phi(\frac{l - np + 0{,}5}{\sqrt{npq}}) - \Phi(\frac{k - np - 0{,}5}{\sqrt{npq}})}}$$

Stellt man keine zu großen Ansprüche an die Genauigkeit, dann reicht aus

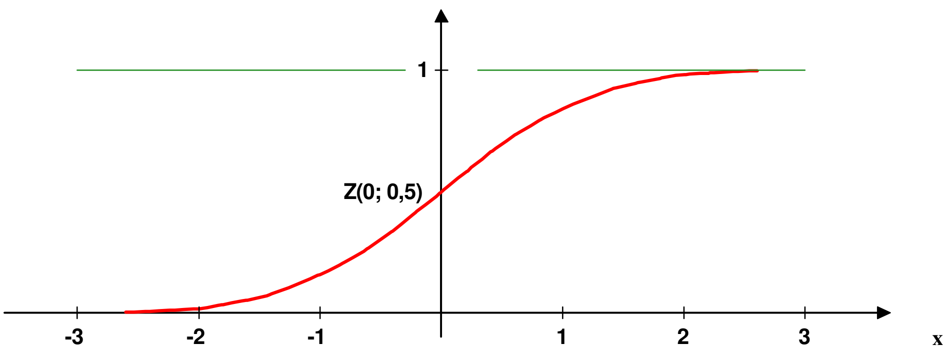
$$\boxed{\mathbf{P(k \le X \le l) \approx \Phi((\frac{l - np}{\sqrt{npq}}) - \Phi((\frac{k - np}{\sqrt{npq}})}}$$

Spezialfall : $l = k$ (**integrale Näherungsformel bzw. Laplace - Näherung**)

$$\boxed{\mathbf{P(X = k) \approx \Phi(\frac{k - np + 0{,}5}{\sqrt{npq}}) - \Phi(\frac{k - np - 0{,}5}{\sqrt{npq}})}}$$

Bemerkungen und Beispiele :

1. Die Werte der Funktion Φ lassen sich nur numerisch berechnen. Als Graph ergibt sich



2. Es gilt $\lim\limits_{x \to -\infty} \Phi(x) = 0$ und $\lim\limits_{x \to \infty} \Phi(x) = 1$

3. Der Graph von Φ ist punktsymmetrisch zu $Z\left(0 \,\middle|\, \frac{1}{2}\right)$. Also ist

$$\Phi(x) - \frac{1}{2} = \frac{1}{2} - \Phi(-x) \quad \Leftrightarrow \quad \boxed{\mathbf{\Phi(-x) = 1 - \Phi(x)}}$$

Daher findet man nur für positive x die Werte von $\Phi(x)$ in der Stochastiktabelle.

So ist z.B. $\Phi(-1{,}23) = 1 - \Phi(1{,}23) \approx 1 - 0{,}88877 = 0{,}11123$

Die Punktsmmetrie von Φ folgt aus der Achsensymmetrie von φ.

4. Φ ist streng monoton steigend und damit umkehrbar. Man findet Werte der Umkehrfunktion in der Stochastiktabelle.

   So ist z.B. $\Phi^{-1}(0{,}0{,}95) \approx 1{,}6449$ und $\Phi^{-1}(0{,}975) \approx 1{,}9600$

5. Für die Verteilungsfunktion F gilt :

$$\boxed{\mathbf{F(x) = P(X \le x) \approx \Phi\left(\frac{x - np + 0{,}5}{\sqrt{npq}}\right) - \Phi\left(\frac{0 - np - 0{,}5}{\sqrt{npq}}\right) \approx \left(\frac{x - np + 0{,}5}{\sqrt{npq}}\right)}}$$

Aufgabentypen :

Berechnen von Wahrscheinlichkeiten

Ein L-Würfel wird 600mal geworfen. Berechnen Sie näherungsweise die Wahrscheinlichkeit, dass

a) mindestens 95mal und höchstens 110mal

b) höchstens 105 mal

c) mindestens 102mal

d) genau 99mal

eine Sechs fällt.

Lösung :

X : Anzahl der gewürfelten Sechsen

Erwartungswert : $E(X) = \mu = np = 600 \cdot \frac{1}{6} = 100$

Standardabweichung : $\sigma = \sqrt{Var(X)} = \sqrt{npq} = \sqrt{600 \cdot \frac{5}{6} \cdot \frac{1}{6}} = 5\sqrt{\frac{10}{3}}$

a) $P(95 \le X \le 110) \approx \phi(\frac{110-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) - \phi(\frac{95-100-\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) \approx \Phi(1{,}15) - \Phi(-0{,}60) =$

$\approx \Phi(1{,}15) - \left[1 - \Phi(0{,}60)\right] \approx\approx 0{,}93943 - 1 + 0{,}72575 \approx 66{,}5\%$

bzw.

$P(95 \le X \le 110) = P(X \le 110) - P(X \le 94) = F(110) - F(94) =$

$= \Phi(\frac{110-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) - \Phi(\frac{94-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}})$

mit dem gleichen Ergebnis.

b) $P(0 \leq X \leq 105) = \Phi(\frac{105-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) - \Phi(\frac{0-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) \approx$

$\approx \Phi(0{,}60) - \underbrace{\Phi(-10{,}90)}_{\approx 0} \approx 0{,}725575 \approx 72{,}6\%$

c) $P(X \geq 102) = 1 - P(X \leq 102) \approx 1 - \Phi(\frac{102-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) \approx$

$\approx 1 - \Phi(0{,}27) \approx 1 - 0{,}60642 \approx 39{,}4\%$

d) $P(X = 99) \approx \Phi((\frac{99-100+\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) - \Phi((\frac{99-100-\frac{1}{2}}{5\sqrt{\frac{10}{3}}}) \approx$

$\approx \Phi(-0{,}05) - \Phi((-0{,}16) = \Phi((0{,}16) - \Phi((0{,}05) \approx 4{,}3\%$

Bestimmen von Intervallen

Eine Laplace-Münze wird 400mal geworfen.

In welchem möglichst kleinen Intervall symmetrisch zum Erwartungswert liegt mit mindestens 95% Wahrscheinlichkeit die Anzahl der Adler ?

Lösung :

X : Anzahl der Adler

Erwartungswert : $E(X) = \frac{1}{2} \cdot 400 = 200$

Standardabweichung : $\sigma = \sqrt{400 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 10$

Ansatz für das Intervall : $I = \left[200-k;\, 200+k\right]$

Bedingung : $P(200-k \leq X \leq 200+k) \geq 0{,}95$

$\Leftrightarrow \Phi(\frac{200+k-200+\frac{1}{2}}{10}) - \Phi(\frac{200-k-200-\frac{1}{2}}{10}) \geq 0{,}95$

$$\Leftrightarrow \Phi\left(\frac{k+\frac{1}{2}}{10}\right) - \Phi\left(\frac{-k-\frac{1}{2}}{10}\right) \geq 0{,}95$$

$$\Leftrightarrow \Phi\left(\frac{k+\frac{1}{2}}{10}\right) - \left[1-\Phi\left(\frac{k+\frac{1}{2}}{10}\right)\right] \geq 0{,}95$$

$$\Leftrightarrow 2\Phi\left(\frac{k+\frac{1}{2}}{10}\right) \geq 1{,}95$$

$$\Leftrightarrow \Phi\left(\frac{k+\frac{1}{2}}{10}\right) \geq 0{,}975 \quad \Big| \circ \, \phi^{-1}$$

$$\Leftrightarrow \frac{k+\frac{1}{2}}{10} \geq 1{,}96 \quad \Leftrightarrow k \geq 19{,}1 \quad k_{min} = 20$$

Mit mindestens 95% Wahrscheinlichkeit liegt die Anzahl der Adler im Intervall $\left[180;\, 220\right]$.

Bestimmen der Versuchsanzahl

Aufgabe :

In einer Urne beträgt der Anteil roter Kugeln 36%. Wie viele Kugeln muss man mindestens mit Zurücklegen ziehen, damit man mit mindestens 90%-iger Wahrscheinlichkeit mehr als 100 rote Kugeln darunter sind ?

Lösung :

Bedingung : $P(X > 100) \geq 0{,}9 \;\Leftrightarrow\; 1 - P(X \leq 100) \geq 0{,}90 \;\Leftrightarrow\; P(X \leq 100) \leq 0{,}1$

$$\Phi\left(\frac{100 - 0{,}36n + 0{,}5}{\sqrt{n \cdot 0{,}36 \cdot 0{,}64}}\right) \leq 0{,}1$$

$$\frac{100{,}5 - 0{,}36n}{0{,}48\sqrt{n}} \leq -1{,}2816 \;\Leftrightarrow\; 100{,}5 - 0{,}36n \leq -0{,}615168\sqrt{n} \;\Leftrightarrow\; 0{,}36n - 0{,}615168\sqrt{n} - 100{,}5 \geq 0$$

Substitution $u = \sqrt{n}$ und Übergang zur Gleichung :

$$0{,}36u^2 - 0{,}615168u - 100{,}5 = 0 \;\Leftrightarrow u \approx 17{,}6 \vee \left[u \approx -15{,}9\right]$$

Rücksubstitution und Rundung : Man muss mindestens 321 ziehen.

Bestimmen der Trefferwahrscheinlicheit

Wie hoch muss der Anteil roter Kugeln in einer Urne mindestens sei, damit man beim Ziehen von 400 Kugeln mit Zurücklegen mit mindestens 90% Wahrscheinlichkeit mindestens 100 rote Kugeln erhält ?

Lösung :

$P(X \geq 100) \geq 0{,}9 \;\Leftrightarrow\; 1 - P(X \leq 99) \geq 0{,}90 \;\Leftrightarrow\; P(X \leq 99) \leq 0{,}1$

$$\Phi\left(\frac{99 - 400p + 0{,}5}{\sqrt{400pq}}\right) \leq 0{,}1$$

$$\Leftrightarrow \frac{99{,}5 - 400p}{20\sqrt{p(1-p)}} \leq -1{,}2816 \;\Leftrightarrow\; 99{,}5 - 400p \leq -25{,}632\sqrt{p(1-p)} \;\Leftrightarrow$$

$$160656{,}9994p^2 - 80256{,}99942p - 9900{,}25 \geq 0$$

Daraus ergibt sich, dass der Anteil mindestens 27,8% betragen muss.

Normalverteilte Größen

Die mittlere Lebensdauer eines Motors beträgt 200000 km mit einer Standardabweichung von 40000 km.

Wie groß ist die Wahrscheinlichkeit, dass ein Motor eine Lebensdauer von mindestens 250000 km hat ?

Lösung :

$$P(X \geq 250000) = 1 - P(X \leq 250000) = 1 - \Phi\left(\frac{250000 - 200000}{40000}\right) = 1 - \Phi(1{,}25) \approx 11{,}6\%$$

Definition :

Hat eine Zufallsgröße X mit $E(X) = \mu$ und $Var(X) = \sigma^2$ die Verteilungsfunktion

$$\mathbf{F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)},$$

dann heißt sie **normalverteilt** nach $N(\mu; \sigma)$.