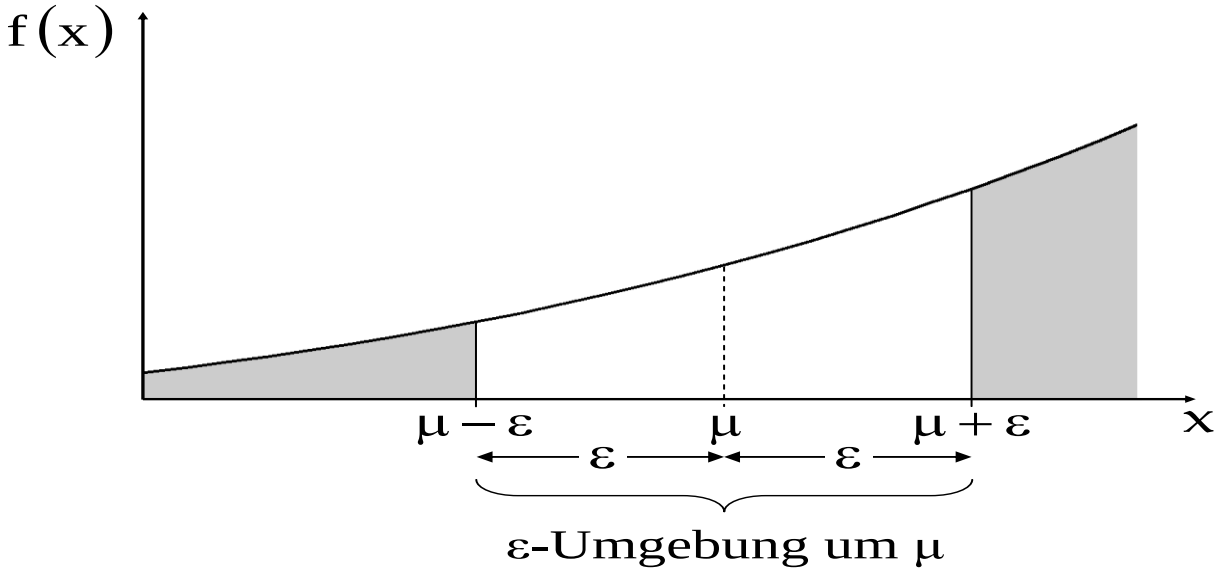
6. Grenzwertsätze

6.1 Tschebyscheffsche Ungleichung

Sofern für eine Zufallsvariable X die Verteilung bekannt ist, lässt sich die Wahrscheinlichkeit dafür bestimmen, dass X in einem bestimmten Intervall liegt. Wie ist aber eine solche Wahrscheinlichkeit zu ermitteln, wenn der Verteilungstyp nicht bekannt ist?

Gefragt ist nach der Wahrscheinlichkeit, dass eine beliebig verteilte Zufallsvariable innerhalb oder außerhalb einer ε-Umgebung um den Erwartungswert μ liegt.



Mit der Tschebyscheffschen Ungleichung kann eine Höchstwahrscheinlichkeit dafür bestimmt werden, dass die Zufallsvariable X mindestens um ε vom Erwartungswert μ abweicht.

Die Wahrscheinlichkeit, dass eine beliebig verteilte Zufallsvariable X mit dem Erwartungswert μ und der Varianz σ^2 um mindestens ε von μ abweicht, ist höchstens gleich σ^2/ε^2:

$$P\left(\left|X - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}$$

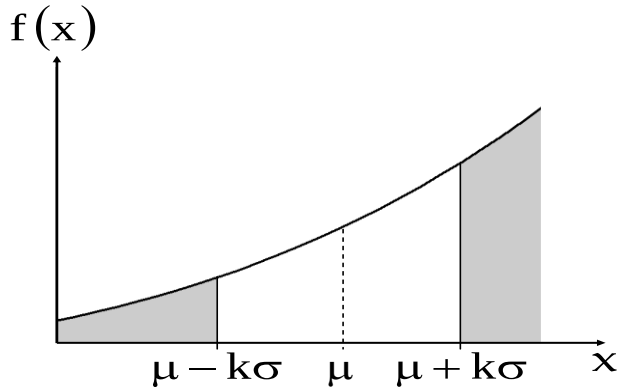
Aufgrund der Wahrscheinlichkeit des Komplementärereignisses ergibt sich eine gleichwertige Formulierung der Tschebyscheffschen Ungleichung, die eine Mindestwahrscheinlichkeit angibt:

$$P\left(\left|X - \mu\right| < \varepsilon\right) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Falls speziell ε = k·σ, folgt:

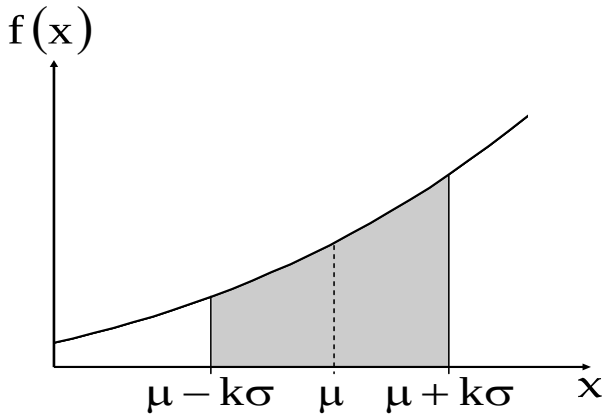
$$P\left(\left|X - \mu\right| \geq k\sigma\right) \leq \frac{1}{k^2}$$

Die Höchstwahrscheinlichkeit, dass eine Zufallsvariable X um mindestens dem k-fachen ihrer Standardabweichung vom Erwartungswert abweicht, beträgt $1/k^2$.



Wird die Tschebyscheffsche Ungleichung für Mindestwahrscheinlichkeiten formuliert, ergibt sich:

$$P\left(\left|X - \mu\right| < k\sigma\right) \geq 1 - \frac{1}{k^2}$$



Auswirkung von k auf die Höhe der Mindestwahrscheinlichkeiten

Berechnet wird die Mindestwahrscheinlichkeit, dass eine beliebig verteilte Zufallsvariable X um nicht mehr als das k = 1, 2, 3-fache ihrer Standardabweichung von ihrem Erwartungswert abweicht:

k	Intervall	Mindestwahrscheinlichkeit
1	$(\mu - \sigma, \mu + \sigma)$	$P\left(\lvert X - \mu \rvert < \sigma\right) \geq 1 - \frac{1}{1^2} = 1 - 1 = 0$
2	$(\mu - 2\sigma, \mu + 2\sigma)$	$P\left(\lvert X - \mu \rvert < 2\sigma\right) \geq 1 - \frac{1}{2^2} = 1 - 0{,}25 = 0{,}75$
3	$(\mu - 3\sigma, \mu + 3\sigma)$	$P\left(\lvert X - \mu \rvert < 3\sigma\right) \geq 1 - \frac{1}{3^2} = 1 - 0{,}111 = 0{,}889$

Für k = 1 ergibt sich kein Informationsgewinn, da Wahrscheinlichkeiten stets nicht negativ sind. Für k > 1 nehmen die Mindestwahrscheinlichkeiten zu, wenn k steigt.

Beispiel:

Der Benzinpreis hängt bei gegebenem Nachfrageverhalten der Autofahrer von Einflussgrößen wie z.B. dem Rohölpreis und dem Bestand an Kraftfahrzeugen ab. Bei einer Mineralölfirma rechnet man für eine Planungsperiode mit einem durchschnittlichen Benzinpreis von 1,70 € pro Liter. Da Preisschwankungen nicht ausgeschlossen werden können, wird von einer Standardabweichung in Höhe von 0,12 € ausgegangen.

Wie groß ist mindestens die Wahrscheinlichkeit, dass der Benzinpreis (= Zufalls-variable X) innerhalb eines 2σ-Bereichs um den Erwartungswert liegt?

Es ist ε = 2σ. Mit der Tschebyscheffschen Ungleichung erhält man für k = 2 folgende Mindestwahrscheinlichkeit:

$$P\left(\left|X - \mu\right| < 2\sigma\right) \geq 1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 1 - 0{,}25 = 0{,}75$$

Wie groß ist höchstens die Wahrscheinlichkeit, dass der Benzinpreis um mehr als 0,15 € vom Erwartungswert abweicht?

Hier gilt ε = 0,15. Mit der Tschebyscheffschen Ungleichung ergibt sich daraus folgende Höchstwahrscheinlichkeit:

$$P\left(\left|X - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2} = \frac{0{,}12^2}{0{,}15^2} = \frac{0{,}014}{0{,}023} = 0{,}609$$

6.2 Gesetze der großen Zahlen

Schwaches Gesetz der großen Zahlen von Tschebyscheff

Ausgangspunkt ist eine Zufallsvariable X mit dem Erwartungswert E(X) = μ und der Varianz Var(X) = σ^2. Der zugrundeliegende Zufallsvorgang wird nun unabhängig n-mal wiederholt.

X_i sei die Zufallsvariable, die angibt, welchen Wert X bei der i-ten Durchführung des Zufallsvorgangs (i = 1, 2, …, n) annimmt.

Der nach der i-ten Durchführung beobachtete Wert x_i (i = 1, 2, … ,n) ist eine Realisation der Zufallsvariablen X_i

Die Zufallsvariablen X_i sind dann nicht nur unabhängig, sondern auch identisch verteilt (d.h. mit einer gleichen Wahrscheinlichkeits- bzw. Dichtefunktion). Insbesondere haben die Zufallsvariablen X_1, X_2, …, X_n auch denselben Erwartungswert und dieselbe Varianz:

$$E(X_i) = \mu$$

$$Var(X_i) = \sigma^2$$

für alle i = 1, 2, ... ,n

Arithmetisches Mittel der Zufallsvariablen X_i	Nach den n Wiederholungen beobachtetes arithmetisches Mittel
$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$

Erwartungswert von $\bar{X}_n$:

$$E\left(\bar{X}_n\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) = \frac{1}{n} n\mu = \mu$$

Der Erwartungswert der Zufallsvariablen $\bar{X}_n$ ist gleich dem Erwartungswert der Zufallsvariablen X.

Varianz von $\bar{X}_n$:

$$Var\left(\bar{X}_n\right) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left(X_i\right) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Die Varianz von $\bar{X}_n$ ist um den Faktor 1/n kleiner ist als die Varianz der Zufalls-variablen X.

Das schwache Gesetz der großen Zahlen von Tschebyscheff besagt, dass die Wahrscheinlichkeit einer Abweichung des arithmetischen Mittels $\overline{X}_n$ vom Erwartungswert μ von weniger als einem beliebigen kleinen Wert ε (ε > 0) mit wachsendem n gegen 1 strebt:

$$\lim_{n\to\infty} P\left(\left|\overline{X}_n - \mu\right| < \varepsilon\right) = 1$$

Beweis:

$$P\left(\left|\overline{X}_n - \mu\right| < \varepsilon\right) \geq 1 - \frac{Var\left(\overline{X}_n\right)}{\varepsilon^2}$$

$$P\left(\left|\overline{X}_n - \mu\right| < \varepsilon\right) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}$$

Mit $\lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$ folgt für $n\to\infty$ hieraus unmittelbar dieses Gesetz der großen Zahlen.

Die Folge von Zufallsvariablen $\overline{X}_n$ konvergiert somit stochastisch gegen μ. Diese stochastische Konvergenz bedeutet, dass es bei n gegen unendlich fast sicher ist, dass $\overline{X}_n$ Werte in einer ε-Umgebung um μ annimmt. Bei großem n liegt $\overline{X}_n$ daher mit hoher Wahrscheinlichkeit in dem beliebig kleinen Intervall (μ - ε, μ + ε). Man kann also mit einer hohen Wahrscheinlichkeit den Erwartungswert μ der Zufallsvariablen X durch eine konkrete Realisierung der Zufallsvariablen $\overline{X}_n$ abschätzen, sofern n genügend groß ist.

Schwaches Gesetz der großen Zahlen von Bernoulli

Hier wird der Spezialfall betrachtet, dass die X_i unabhängige mit dem Parameter p Bernoulli-verteilte Zufallsvariablen sind. Die Summe Y_n von n dieser Zufallsvariablen ist dann binomialverteilt mit den Parametern n und p:

$$Y_n = \sum_{i=1}^{n} X_i$$

Betrachtet wird nun:

$$\overline{P}_n = \frac{1}{n} Y_n$$

Erwartungswert von $\overline{P}_n$:

$$E\left(\overline{P}_n\right) = E\left(\frac{1}{n} Y_n\right) = \frac{1}{n} E\left(Y_n\right) = \frac{1}{n} np = p$$

Varianz von $\overline{P}_n$:

$$Var\left(\overline{P}_n\right) = Var\left(\frac{1}{n} Y_n\right) = \frac{1}{n^2} Var\left(Y_n\right) = \frac{1}{n^2} np\left(1 - p\right) = \frac{p\left(1 - p\right)}{n}$$

Analog zum schwachen Gesetz der großen Zahlen von Tschebyscheff gilt für das schwache Gesetz der großen Zahlen von Bernoulli:

$$\lim_{n \to \infty} P\left(\left|\overline{P}_n - p\right| < \varepsilon\right) = 1$$

6.3 Zentrale Grenzwertsätze

Zentrale Grenzwertsätze machen eine Aussage über das Grenzverhalten einer Folge von Verteilungsfunktionen F_{X_n}, die zu einer Folge von Zufallsvariablen X_n gehört. Dabei gibt es unterschiedliche Verallgemeinerungsgrade und zahlreiche Varianten. Im Kern erhält man das Ergebnis, dass bei Zufallsvariablen, die als Summen oder Durchschnitte interpretiert werden können, bei großer Anzahl von Wiederholungen eines Zufallsvorgangs die Normalverteilung anwendbar ist.

Grenzwertsatz von Lindeberg und Lévy

Beim Grenzwertsatz von Lindeberg und Lévy setzt man keinen speziellen Verteilungstyp voraus. Betrachtet wird eine Summe Y_n von n unabhängigen identisch verteilten Zufallsvariablen X_i mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$.

$$Y_n = \sum_{i=1}^{n} X_i$$

Erwartungswert von Y_n	Varianz von Y_n
$E(Y_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n\mu$	$Var(Y_n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = n\sigma^2$

Die Zufallsvariable Y_n wird nun standardisiert:

$$Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

Daraus folgt:

$$Z_n \overset{a}{\sim} N(0;\ 1)$$

Der Grenzwertsatz von Lindeberg und Lévy besagt, dass die standardisierte Summe Z_n von unabhängigen identisch verteilten Zufallsvariablen asymptotisch standardnormalverteilt ist.

Bei endlichem n wird für n > 30 (Faustregel) näherungsweise die Standard-normalverteilung verwendet.:

Daraus folgt, dass die Summe unabhängiger und identisch verteilter Zufallsvariablen X_i für große n (d.h. für n > 30) näherungsweise normalverteilt ist:

$$Y_n \overset{appr}{\sim} N(n\mu;\ n\sigma^2)$$

Beispiel:

Eine Einzelhandelskette plant, in einer Region mit 1 Million Haushalten ein Filialnetz zu errichten. Das Marktvolumen beträgt gegenwärtig 150 Millionen € pro Quartal, was einem durchschnittlichen Umsatz von 150 € je Haushalt entspricht. Die Standardabweichung der Haushaltsausgaben für diese Einzelhandelskette pro Quartal liegt ebenfalls bei 150 €. Eine Testfiliale, die in angemessener Zeit für 900 Haushalte erreichbar ist, hat in einem Quartal einen Gesamtumsatz von 144000 € erzielt.

Die Marktforschungsabteilung wird damit beauftragt, die Wahrscheinlichkeit für einen in dieser Höhe erzielten oder größeren Umsatz unter unveränderten Marktbedingungen zu bestimmen.

Die Zufallsvariable X_i bezeichnet die Ausgaben des i-ten Haushalts in €. Daraus folgt $E(X_i) = \mu = 150$ und $\sigma = 150$ für i = 1, 2, ,..., 900. Es wird angenommen, dass die X_i unabhängige und identisch verteilte Zufallsvariablen sind. Dies bedeutet, dass sich die Ausgaben unterschiedlicher Haushalte nicht gegenseitig beeinflussen. Für die Summe der gesamten in der Testfiliale erzielten Ausgaben der Haushalte gilt dann:

$$Y_n = \sum_{i=1}^{900} X_i$$

Für den Erwartungswert und die Standardabweichung von Y_n erhält man:

$E(Y_n) = n \cdot \mu = 900 \cdot 150 = 135000$

$$\sqrt{\mathrm{Var}(Y_n)} = \sigma\sqrt{n} = 150 \cdot \sqrt{900} = 150 \cdot 30 = 4500$$

Die Summe der Ausgaben 144000 € für alle n = 900 Haushalte ist die konkrete Realisation der Zufallsvariablen Y_n. Gesucht wird nun die Wahrscheinlichkeit, dass Y_n Werte annimmt, die größer oder gleich 144000 sind:

$P(Y_n \geq 144000)$

Daraus folgt:

$P(Y_n \geq 144000) = 1 - P(Y_n < 144000)$

Da hier die Voraussetzungen des Zentralen Grenzwertsatzes von Lindeberg und Lévy erfüllt sind (n = 900 > 30), lässt sich die gesuchte Wahrscheinlichkeit approximativ über die Verteilungsfunktion der Standardnormalverteilung bestimmen:

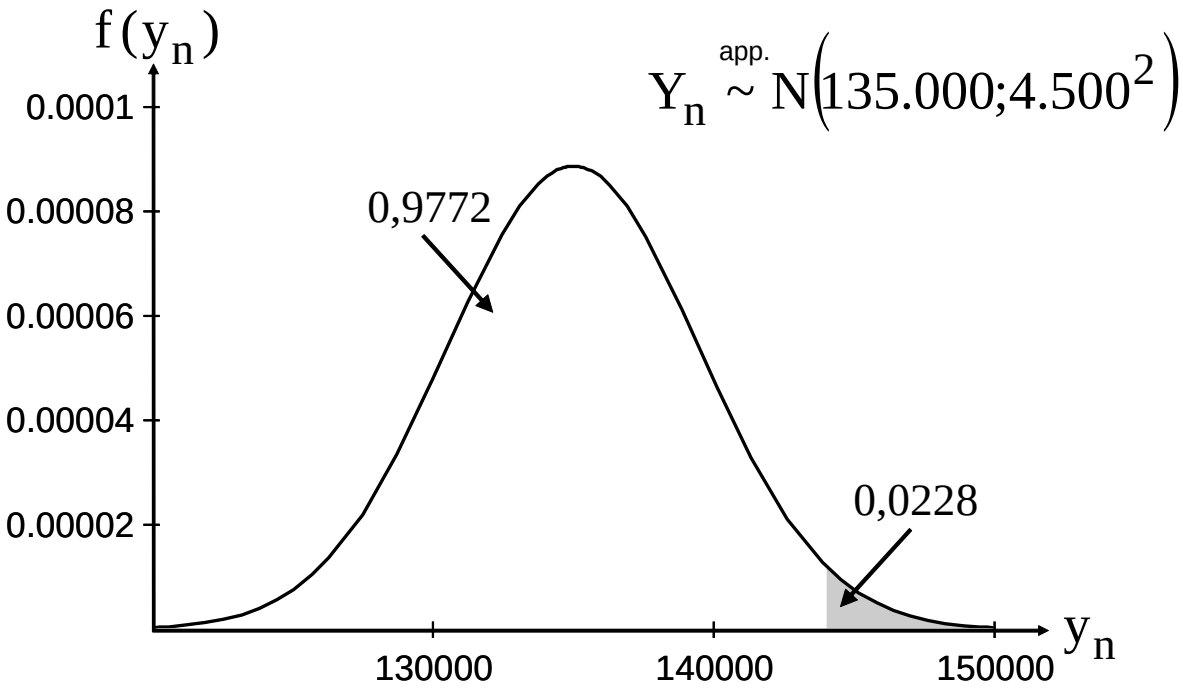
$$P(Y_n < 144000) \approx \Phi\left(\frac{144000 - 135000}{4500}\right) = \Phi(2) = 0{,}9772$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817

Damit ergibt sich:

$$P(Y_n \geq 144000) \approx 1 - \Phi(2) = 1 - 0{,}9772 = 0{,}023$$

Die Wahrscheinlichkeit, bei unveränderten Marktbedingungen einen Umsatz von mindestens 144000 € zu erzielen, liegt somit bei 2,3%.



Betrachtet wird nun das arithmetische Mittel $\bar{X}_n$ als Lineartransformation der Summenvariablen Y_n (a = 0, b = 1/n):

$$\bar{X}_n = \frac{1}{n} Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Bei der Lineartransformation bleibt die approximative Normalverteilung erhalten. Auch $\bar{X}_n$ wird nun standardisiert:

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}$$

Nach dem Grenzwertsatz von Lindeberg und Lévy gilt dann, dass Z_n asymptotisch standardnormalverteilt ist:

$$Z_n \overset{a}{\sim} N(0; 1)$$

Daraus folgt, dass das arithmetische Mittel $\bar{X}_n$ für große n (d.h. für n > 30) näherungsweise normalverteilt ist:

$$\bar{X}_n \overset{appr}{\sim} N\left(\mu; \frac{\sigma^2}{n}\right)$$

Grenzwertsatz von de Moivre und Laplace

Hier wird nun der Spezialfall betrachtet, dass die X_i unabhängige Bernoulli-verteilte Zufallsvariablen sind.

Die Summenvariable $Y_n = \sum_{i=1}^{n} X_i$ ist daher binomialverteilt.

Die Zufallsvariable Y_n wird erneut standardisiert:

$$Z_n = \frac{Y_n - np}{\sqrt{np(1-p)}}$$

Daraus folgt:

$$Z_n \overset{a}{\sim} N(0; 1)$$

Der Grenzwertsatz von de Moivre und Laplace besagt, dass die standardisierte Summe Z_n von unabhängigen identisch Bernoulli-verteilten Zufallsvariablen asymptotisch standardnormalverteilt ist.

Daraus folgt, dass die Summe unabhängiger Bernoulli-verteilter Zufallsvariablen X_i, d.h. eine binomialverteilte Zufallsvariable Y_n, für große n näherungsweise normalverteilt ist:

$$Y_n \overset{appr}{\sim} N\left[np; np(1-p)\right]$$

Für die Praxis benötigen wir eine Information, wie groß n mindestens sein muss, damit diese Approximation anwendbar ist. Verwendet wird häufig folgende Faustregel:

$$n > \frac{9}{p(1-p)}$$

Demnach muss n um so größer sein, je asymmetrischer die Binomialverteilung ist:

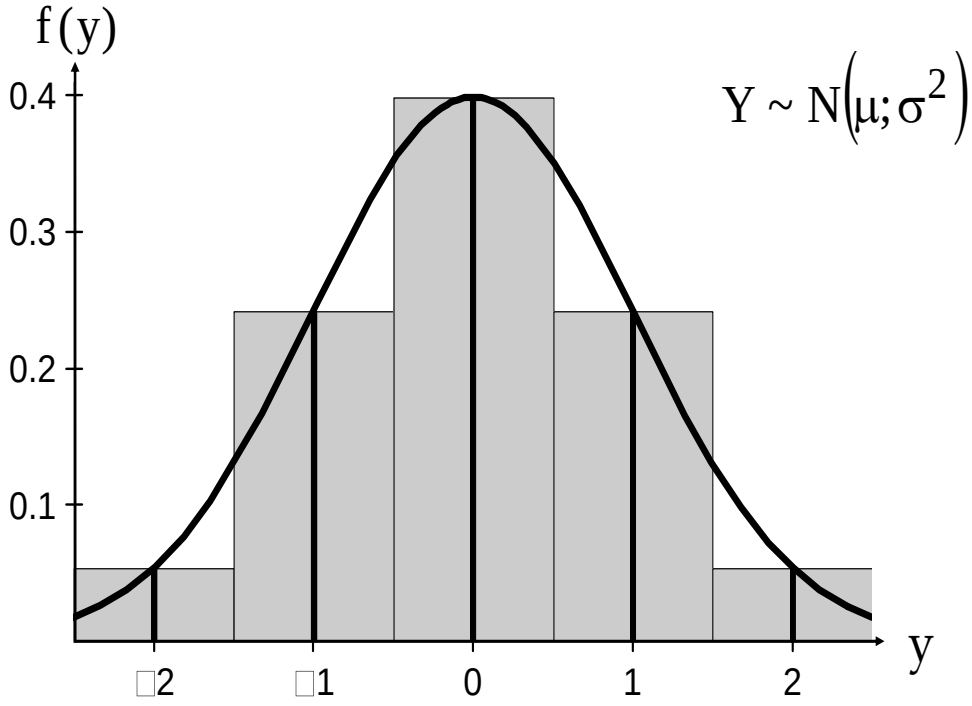
- Bei Symmetrie, d.h. $p = 0{,}5$: $n > 36$
- Bei merklicher Asymmetrie, z.B. $p = 0{,}2$: $n > 56$
- Bei starker Asymmetrie, z.B. $p = 0{,}1$: $n > 100$

Stetigkeitskorrektur

Bei der Approximation der Binomial- durch die Normalverteilung nimmt man häufig noch eine Stetigkeitskorrektur vor, d.h. im Argument der Verteilungsfunktionen der Normalverteilung wird ±0,5 berücksichtigt. Je größer n ist, desto weniger wirkt sich die Stetigkeitskorrektur auf die Berechnung der Wahrscheinlichkeiten aus.

Eine binomialverteilte Zufallsvariable Y kann nur die Werte $y = 0, 1, ..., n$ mit den Wahrscheinlichkeiten $P(Y = y)$ annehmen. Deshalb konstruiert man Säulen mit einer Breite von 1 um die Stäbe.

Visualisierung der Stetigkeitskorrektur



Wahrscheinlichkeit, dass eine binomialverteilte Zufallsvariable Y einen bestimmten Wert a annimmt:

$$P(Y = a) \approx F(a + 0{,}5) - F(a - 0{,}5)$$

F: Verteilungsfunktion der Normalverteilung

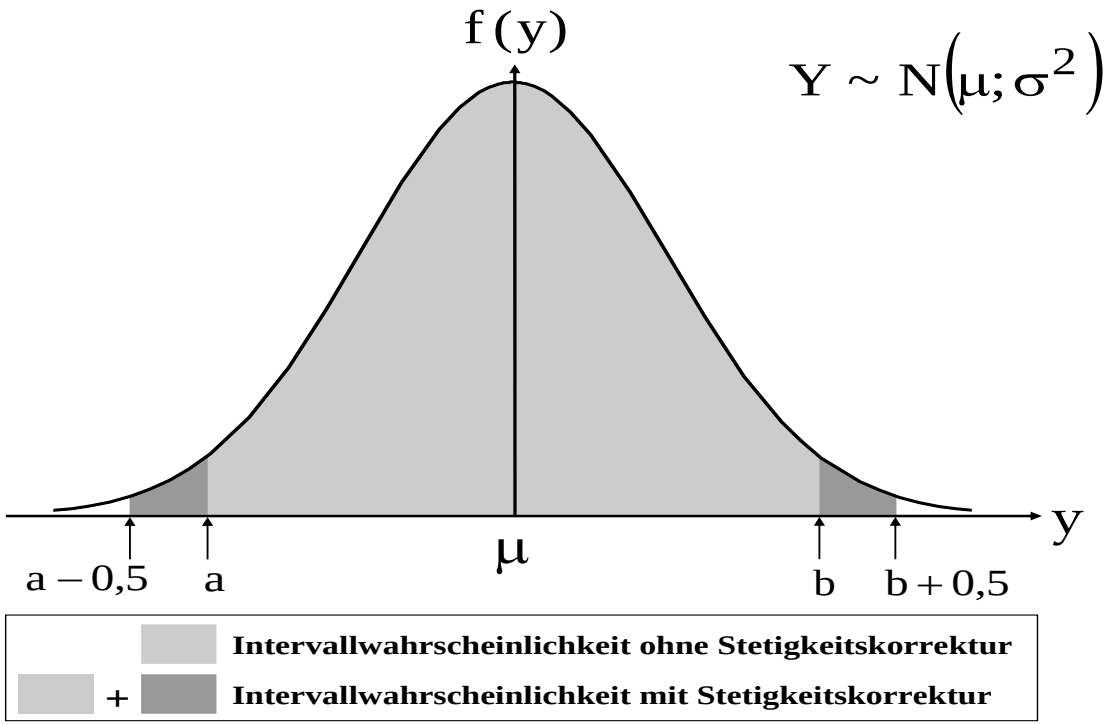
(Fläche des Rechtecks über dem Intervall [a-0,5; a+0,5])

Wahrscheinlichkeit, dass eine binomialverteilte Zufallsvariable zwischen den Grenzen a und b liegt (mit Stetigkeitskorrektur):

$$P(a \leq Y \leq b) \approx F(b + 0{,}5) - F(a - 0{,}5)$$

Da F(y) die Verteilungsfunktion einer Normalverteilung mit den Parametern E(Y) = n·p und Var(Y) = n·p·(1-p) ist, gilt für die Intervallwahrscheinlichkeit:

$$P(a \leq Y \leq b) \approx \Phi\left(\frac{b + 0{,}5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - 0{,}5 - np}{\sqrt{np(1-p)}}\right)$$



Beispiel:

Ein Betrieb liefert Glühlampen in Kartons zu je 1000 Stück. Aus früheren Untersuchungen ist bekannt, dass der Betrieb im Durchschnitt 3% Ausschuss produziert.

Wie groß ist die Wahrscheinlichkeit, dass sich in einem zufällig ausgewählten Karton zwischen 20 und 40 defekte Glühlampen befinden?

Es sei A das Ereignis, dass eine entnommene Glühlampe defekt ist. Wenn A eintritt, d.h. die entnommene Glühbirne defekt ist, nimmt die Zufallsvariable X den Wert 1 an, sonst 0. Die Zufallsvariable Y bezeichnet die Anzahl der defekten Glühlampen in dem zufällig ausgewählten Karton:

$$Y = \sum_{i=1}^{1000} X_i$$

Y ist binomialverteilt mit den Parametern n = 1000 und p = 0,03. Die gesuchte Wahrscheinlichkeit dafür, dass Y in das Intervall [20, 40] fällt, ergibt sich dann folgendermaßen:

$$P\left(20 \le Y \le 40\right) = \sum_{y=20}^{40} \binom{1000}{y} 0{,}03^{y}\, 0{,}97^{1000-y}$$

Eine Berechnung der Wahrscheinlichkeit mit dieser Formel ist jedoch sehr umständlich. Da n = 1000 aber ausreichend groß ist, kann eine Approximation mit Hilfe der Normalverteilung durchgeführt werden:

$$n > \frac{9}{p(1-p)} = \frac{9}{0{,}03 \cdot (1-0{,}03)} = 309{,}278$$

Mit den Parametern E(Y) = np = 1000·0,03 = 30 und Var(Y) = np(1-p) = 1000·0,03·0,97 = 29,1 ergibt sich:

$$P(20 \le Y \le 40) \approx \Phi\left(\frac{40-30}{\sqrt{29{,}1}}\right) - \Phi\left(\frac{20-30}{\sqrt{29{,}1}}\right) = \Phi(1{,}85) - \Phi(-1{,}85)$$

Aufgrund der Symmetrie der Standardnormalverteilung gilt $\Phi(-1{,}85) = 1 - \Phi(1{,}85)$.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	$\boxed{0{,}9678}$	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767

Damit erhält man die gesuchte Wahrscheinlichkeit:

$$P(20 \le Y \le 40) \approx \Phi(1{,}85) - \Phi(-1{,}85) = \Phi(1{,}85) - \left[1 - \Phi(1{,}85)\right]$$

$$= 0{,}9678 - (1 - 0{,}9678) = 0{,}936$$

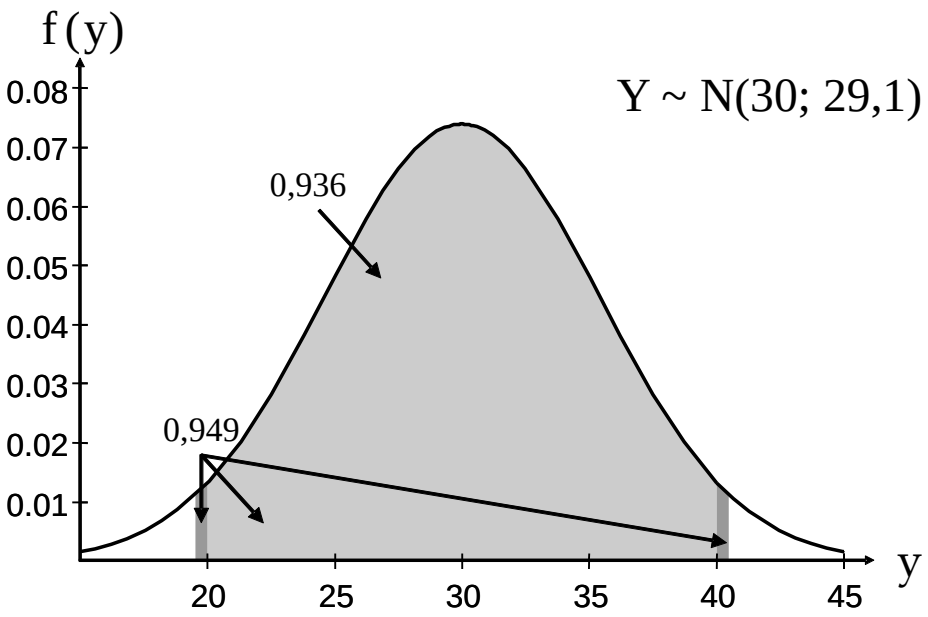
Allerdings wurde hier nicht berücksichtigt, dass bei der Approximation der Binomial- durch die Normalverteilung eine diskrete durch eine stetige Verteilung angenähert wird. Wenn die gesuchte Wahrscheinlichkeit unter Einbeziehung der Stetigkeitskorrektur berechnet wird, ergibt sich:

$$P(20 \le Y \le 40) \approx \Phi\left(\frac{40 + 0{,}5 - 30}{\sqrt{29{,}1}}\right) - \Phi\left(\frac{20 - 0{,}5 - 30}{\sqrt{29{,}1}}\right)$$

$$= \Phi(1{,}95) - \Phi(-1{,}95) = \Phi(1{,}95) - \left[1 - \Phi(1{,}95)\right]$$

$$= 0{,}9744 - (1 - 0{,}9744) = 0{,}949$$

Wenn also die Stetigkeitskorrektur vorgenommen wird, erhöht sich die Wahrscheinlichkeit um etwa 1,3 Prozentpunkte.



Betrachtet wird nun das arithmetische Mittel $\overline{P}_n$ als Lineartransformation der Summenvariablen Y_n (a = 0, b = 1/n):

$$\overline{P}_n = \frac{1}{n} Y_n$$

Bei der Lineartransformation bleibt die approximative Normalverteilung erhalten. Auch $\overline{P}_n$ wird nun standardisiert:

$$Z_n = \frac{\overline{P}_n - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Nach dem Grenzwertsatz von de Moivre und Laplace gilt dann, dass Z_n asymptotisch standardnormalverteilt ist:

$$Z_n \overset{a}{\sim} N(0; 1)$$

Daraus folgt, dass $\overline{P}_n$ für große n (d.h. für n > 30) näherungsweise normalverteilt ist:

$$\overline{P}_n \overset{appr}{\sim} N\left(p; \frac{p(1-p)}{n}\right)$$