Normalverteilung

Viele Zufallsprozesse oder andere natürliche Verteilungen folgen annähernd einer Glockenkurve, sind also **normalverteilt**. Eine Zufallsvariable X ist normalverteilt mit einem Erwartungs- bzw. Mittelwert μ und einer Standardabweichung σ — die durchschnittlich zu erwartende Streuung der Messwerte um den Erwartungswert — wenn ihre Wahrscheinlichkeit, einen Wert aus dem Intervall $[x_0; x_1]$ anzunehmen, folgendermaßen ausgedrückt werden kann:

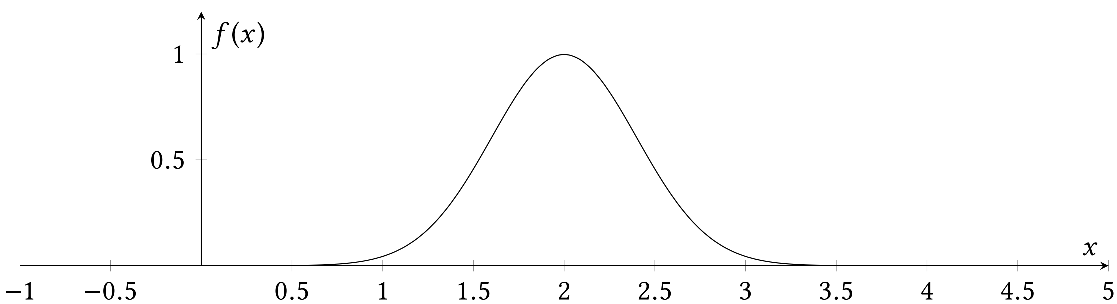
$$P(x_0 \leq X \leq x_1) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{x_0}^{x_1} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Erwartungswert μ	… der durschnittlich zu erwartende Wert der Zufallsvariable X
Standardabweichung σ	… die durschnittlich zu erwartende Abweichung der Zufallsvariable X vom Erwartungswert μ
$X \sim \mathcal{N}(\mu; \sigma)$	… eine normalverteilte Zufallsvariable X mit den Parametern μ und σ

Es ist wichtig anzumerken, dass eine normalverteilte Zufallsvariable nicht diskret, sondern stetig ist. Eine diskrete Zufallsvariable, beispielsweise jene bei der binomialen oder hypergeometrischen Verteilung, kann nur natürliche Zahlenwerte aus einer vordefinierten Teilmenge $\Omega \subset \mathbb{N}$ annehmen, der sogenannten *Ereignismenge*. Gegensätzlich dazu kann eine stetige Zufallsvariable alle möglichen Werte aus der Menge der reellen Zahlen $\mathbb{R}$ annehmen, da sie durch Messung entsteht. Da man zwischen zwei reellen Zahlen stets eine weitere reelle Zahl einfügen kann, ist die Anzahl an möglichen Werten für eine reelle Zahl unendlich. So kann man beispielsweise bei einer stetigen Zufallsvariable, für welche die Grundmenge $\mathbb{G}$ die Menge der reellen Zahlen $\mathbb{R}$ ist, niemals einen Wasserstand, eine Füllmenge oder eine Baumhöhe genau, *auf das Atom* berechnen. Man kann immer nur ein Intervall betrachten, in welchem sich die stetige, reelle Zufallsvariable befindet. Denn die Wahrscheinlichkeit für einen exakten Wert, beispielsweise einen Meter Wasserstand, auf das Atom genau, kein Wasserstoff mehr oder weniger, ist immer gleich null.

Daraus folgt, dass die Wahrscheinlichkeitsberechnung bei der Normalverteilung immer für Intervalle gilt und somit ein Integral ist. Berechnet man die Wahrscheinlichkeit, dass die normalverteilte, stetige Zufallsvariable X einen Wert aus einem Intervall $]x_0; x_1[$ (oder $[x_0; x_1]$) annimmt, so berechnet man das Integral der Normalverteilung in diesem Intervall. Man spricht bei der Normalverteilung also von einer *Wahrscheinlichkeitsdichtefunktion*, bzw. Probability-density-function (PDF).

Eine normalverteilte Zufallsvariable $X \sim \mathcal{N}(2; 0.4)$ ist grafisch folgendermaßen dargestellt (μ bestimmt x-Position, σ bestimmt Breite und Höhe der Kurve):



Standardnormalverteilung

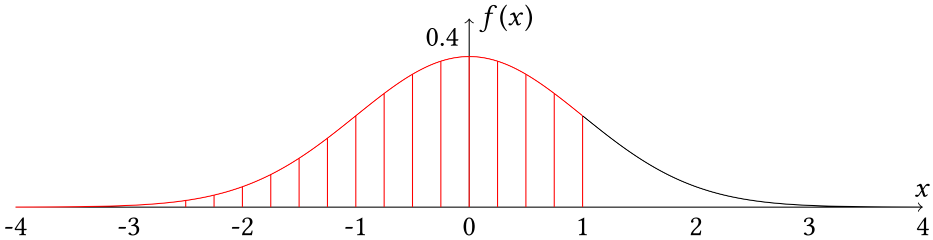
Eine besondere Form der Normalverteilung ist $Z \sim \mathcal{N}(0;1)$, also jene normalverteilte Zufallsvariable Z mit Erwartungswert $\mu = 0$ und Standardabweichung $\sigma = 1$. Ihre Funktion lautet:

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

Die Verteilungsfunktion der Standardnormalverteilung ist jene Funktion $\Phi(z)$, welche für eine beliebige Zufallsvariable $Z \sim \mathcal{N}(0;1)$ das Integral von $-\infty$ bis z berechnet:

$$\Phi(z) = \int_{-\infty}^{z} \varphi(z)\, dz$$

So wäre beispielsweise $\Phi(Z \leq 1) = \int_{-\infty}^{1} \varphi(z)\, dz$:



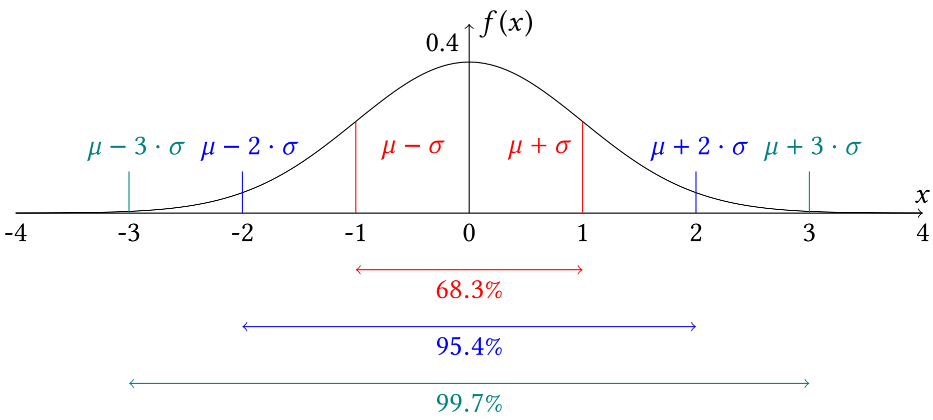
Möchte man die Wahrscheinlichkeit berechnen, dass Z nicht höchstens, sondern mindestens z beträgt, also die Fläche unter der Kurve nach z, von z bis $+\infty$, so wäre dies entweder die Gegenwahrscheinlichkeit zu $\Phi(z)$, also $1 - \Phi(z)$, oder $\Phi(-z)$, da das Intervall über z die selbe Fläche hat wie das Intervall bis $-z$, also das symmetrische Gegenüber auf der anderen Seite von μ bzw. hier der y-Achse. Die Wahrscheinlichkeit, dass Z in einem bestimmten Intervall $]z_1; z_2[$ liegt, ist gleich dem Integral von $\varphi(z)$ in den Grenzen $-\infty$ bis z_2, minus dem Integral von $\varphi(z)$ in den Grenzen $-\infty$ bis z_1. Diese Differenz ist dann der Unterschied zwischen den beiden Flächen, somit das Intervall dazwischen. Formal gesehen:

$$\Phi(z_1 < Z < z_2) = \Phi(z_2) - \Phi(z_1)$$

Letztlich gibt es noch die Funktion $D(z_\varepsilon)$, welche equivalent zu $\Phi(z_\varepsilon) - \Phi(-z_\varepsilon)$ ist. Sie ist also für symmetrische Intervalle $[\mu - \varepsilon; \mu + \varepsilon]$ anwendbar. $D(z_\varepsilon)$ ist ebenso equivalent zu $2 \cdot \Phi(z_\varepsilon) - 1$. Der Beweis für den letzten Ausdruck liegt darin, dass bei einem symmetrischen Intervall mit Parameter ε die untere Grenze gleich weit von μ entfernt ist wie die obere, sprich das Intervall über $\mu + \varepsilon$ ist das selbe wie jenes unter $\mu - \varepsilon$. Da die Wahrscheinlichkeit, dass Z über einem bestimmten z liegt, gleich $1 - \Phi(z)$, kann man diesen Ausdruck für die untere Grenze substituieren:

$$D(z_\varepsilon) = \Phi(z_\varepsilon) - [1 - \Phi(z_\varepsilon)] = 2 \cdot \Phi(z_\varepsilon) - 1$$

Es soll noch gesagt sein, dass die Wahrscheinlichkeitsintervalle der Standardnormalverteilung schon anhand ihrer grafischen Position approximiert werden können. So liegen im Intervall $]\mu - \sigma; \mu + \sigma[$ ungefähr 68.3% Prozent, eine Standardabweichung weiter im Intervall $]\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma[$ schon ungefähr 95.4% und nach drei Standardabweichungen im Intervall $]\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma[$ schon 99.7%.



z-Transformation

Letztlich sei noch anzumerken, dass man jede nicht-standardisierte Normalverteilung zu einer standardisierten transformieren kann. Diesen Vorgang nennt man z-*Transformation*. Dabei wird der Ausdruck $\frac{x-\mu}{\sigma}$ durch die Variable z ersetzt, um somit die nicht-standardisierte Glockenkurve auf der x-Achse so zu verschieben, dass $\mu = 0$ und ihre Breite und Höhe so anzupassen, dass $\sigma = 1$. Dies trägt den Vorteil, dass man nur die Werte der Standardnormalverteilung zu wissen braucht und die Zufallsvariablen aller anderen Normalverteilungen auf diese standardisierten Werte zurückführen kann.

$$\frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\Downarrow$$

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

$$\Downarrow$$

$$\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}$$

Man überlege sich, was die z-Transformation eigentlich ist, bzw. wie sie funktioniert. x ist ein absoluter Wert den die stetige normalverteilte Zufallsvariable X annehmen kann, beispielsweise einen IQ von 110. μ ist in diesem Fall laut Definition des Intelligenz-Quotienten gleich 100 und σ gleich 15. Nun soll dieser normalverteilte Wert 110 mit $\mu = 100$ und $\sigma = 15$ so transformiert werden, dass seine Wahrscheinlichkeit mit der Standardnormalverteilung $\mathcal{N}(\mu = 0; \sigma = 1)$ gefunden werden kann. Die Zufallsvariable Z ist hier die standardnormalverteilte Zufallsvariable, welche um $\mu = 0$ links und rechts schwankt. Zunächst muss der Wert 110 auf den Mittelwert $\mu = 0$ normiert werden. Dazu subtrahiert man den alten Mittelwert $\mu = 100$. Das Ergebnis lautet $x - \mu = 110 - 100 = 10$, schwankt also nun um 0 und ist daher von der Position her passend für die Standardnormalverteilung. Jetzt muss x noch nach der Standardabweichung $\sigma = 1$ normiert werden. Es gilt folgendes Verältnis: $x - \mu$ ist zur alten Standardabweichung $\sigma = 100$, so wie die z-transformierte Variable zur standardisierten Standardabweichung 1:

$$x - \mu : \sigma = z : 1 \Rightarrow \frac{x-\mu}{\sigma} = \frac{z}{1} \Rightarrow z = \frac{x-\mu}{\sigma}$$

Setzt man nun ein, erhält man den z-Wert:

$$z = \frac{110 - 100}{115} = \frac{2}{3}$$

p bestimmen

Um nun die Wahrscheinlichkeit p eines Zufallsereignises bei einer nicht-standardisierten Normalverteilung $\mathcal{N}(\mu; \sigma)$ zu berechnen, muss man zuerst den Parameter z mittels der z-Transformation berechnen und dann den passenden Wahrscheinlichkeitswert aus einer Φ-Tabelle entnehmen.

Beispiel: *Eine Zufallsvariable X ist normalverteilt mit $X \sim \mathcal{N}(\mu = 1.04; \sigma = 0.02)$, berechne die Wahrscheinlichkeit dass die Zufallsvariable einen Wert zwischen 1 und 1.1 annimmt.*

I. Durchführung der z-Transformation für $x_2 = 1.1$: $z_2 = \frac{x_2 - \mu}{\sigma} = \frac{1.1 - 1.04}{0.02} = 3$

II. Durchführung der z-Transformation für $x_1 = 1$: $z_2 = \frac{x_1 - \mu}{\sigma} = \frac{1 - 1.04}{0.02} = -2$

III. Ermittlung der Wahrscheinlichkeit für z_2: $\Phi(3) = 0.9987$

IV. Ermittlung der Wahrscheinlichkeit für z_1: $\Phi(-2) = 0.0228$

V. Berechnung der Differenz und somit p: $\Phi(z_2) - \Phi(z_1) = 0.9987 - 0.0228 = 97.59\%$

Weiters: *Berechne auch die Wahrscheinlichkeit, dass die Zufallsvariable einen Wert höher als 1.02 annimmt.*

I. Durchführung der z-Transformation für $x = 1.02$: $z = \frac{x - \mu}{\sigma} = \frac{1.02 - 1.04}{0.02} = -1$

II. Da es gilt $p(X > x)$ zu ermitteln, also die Fläche unter der Kurve zwischen x und $+\infty$, muss entweder die Gegenwahrscheinlichkeit von $\Phi(z)$ berechnet, also $1 - \Phi(z)$, oder das Vorzeichen von z invertiert werden: $\Phi(-z)$.

III. Ermittlung der Wahrscheinlichkeit für z aus der Φ-Tabelle: $1 - \Phi(-1) = \Phi(1) = 0.8413$

Letztlich: *Berechne zusätzlich noch die Wahrscheinlichkeit, dass die Zufallsvariable einen Wert niedriger als 1 annimmt.*

I. Durchführung der z-Transformation für $x = 1$: $z = \frac{x - \mu}{\sigma} = \frac{1 - 1.04}{0.02} = -2$

II. Da es hier gilt $p(X < x)$ zu ermitteln, also die Fläche unter der Kurve zwischen $-\infty$ und x, kann man $\Phi(z)$ verwenden.

III. Ermittlung der Wahrscheinlichkeit für z aus der Φ-Tabelle: $\Phi(-2) = 0.0288$

Allgemein gilt also:

X **geringer als** x: $P(X < x) = \Phi(z) = \Phi\left(\frac{x - \mu}{\sigma}\right)$

X **größer als** x: $P(X > x) = 1 - \Phi(z) = 1 - \Phi\left(\frac{x - \mu}{\sigma}\right) = \Phi(-z) = \Phi\left[-\left(\frac{x - \mu}{\sigma}\right)\right]$

Symmetrisches Intervall: $P(x_1 < X < x_2) = 2 \cdot \Phi(z_2) - 1 = 2 \cdot \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - 1$

Asymmetrisches Intervall: $P(x_1 < X < x_2) = \Phi(z_2) - \Phi(z_1) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right)$

x bestimmen

Generell erfordert es zur Berechnung des Wertes x einer normalverteilten Zufallsvariable X den Erwartungswert μ, die Standardabweichung σ sowie eine passende Wahrscheinlichkeit p, wo p entweder die Wahrscheinlichkeit $P(X > x)$ oder $P(X < x)$ ist. Auch sei angemerkt, dass in diesem Fall nicht die Φ-Funktion, sondern ihr Invers Φ^{-1} von Interesse ist, da man nicht p für einen bestimmten z-Wert, sondern das passende z für eine Wahrscheinlichkeit p sucht. Dann gilt es, die z-Transformation rückgängig zu machen:

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

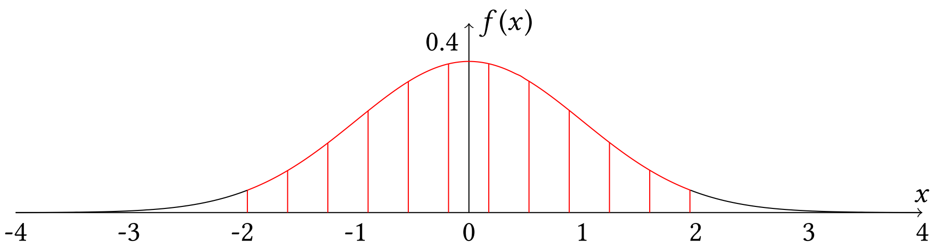
$$\Downarrow$$

$$x = z \cdot \sigma + \mu$$

Beispiel: *Ermittle jenen normalverteilten Wert* $x \sim \mathcal{N}(\mu = 5; \sigma = 1.5)$, *welchen die Zufallsvariable* X *in* 15% *aller Fälle unterschreitet.*

Berechnung von x für $P(Z < x) = 0.15$: $x = z \cdot \sigma + \mu \Rightarrow \Phi^{-1}(0.15) \cdot 1.5 + 5 \Rightarrow -1.035 \cdot 1.5 + 5 = 3.4475$

Beispiel: *Ermittle jenen Wert* ε*, sodass die Zufallsvariable* Z *zu* 95% *im symmetrischen Intervall* $]\mu - \varepsilon; \mu + \varepsilon[$ *liegt, wo* $\mu = 3600$ *und* $\sigma = 200$.



Veranschaulichung

Hierbei muss man zuerst definieren, wie man Z am besten berechnet. In dieser Aufgabenstellung ist im Grunde genommen nur die obere Grenze des symmetrischen Intervalls relevant. Das heißt, man müsste lediglich den x-Wert der oberen Grenze berechnen und von diesem μ subtrahieren, um somit ε zu erhalten. Diesen x-Wert erhalten wir, in dem wir die Wahrscheinlichkeit p in Anbetracht ziehen, dass Z weniger als x ist: $P(Z < \mu + \varepsilon)$. Diese Wahrscheinlichkeit geht jedoch von $-\infty$ bis x, ist also nicht 0.95, sondern 0.95 plus die Wahrscheinlichkeit, dass Z geringer als die untere Grenze ist: $p + P(Z < \mu - \varepsilon)$. Analysiert man die Problemstellung weiter, erkennt man dass diese Wahrscheinlichkeit $P(Z < \mu - \varepsilon)$ gleich $0.5 \cdot (1 - p)$ ist. Dies folgt daraus, dass die Wahrscheinlichkeit außerhalb des Intervalls links und rechts zusammen $1 - p$ ist, einzeln daher die Hälfte davon. Der x-Wert der oberen Grenze $\mu + \varepsilon$ muss also sein:

$$x = z \cdot \sigma + \mu \Rightarrow \Phi^{-1}\left(p + \frac{1 - p}{2}\right) \cdot \sigma + \mu$$

Weitere Beobachtungen zeigen, dass $p + \frac{1-p}{2}$ gleich $\frac{1+p}{2}$ ist. Ebenso ist hier anzumerken, dass die Formel oben dazu dient, den x-Wert der oberen Grenze zu berechnen. Die Aufgabenstellung verlangt jedoch ε, den Abstand dieses x-Wertes der oberen Grenze vom Erwartungswert μ. Es sollte also eigentlich heißen:

$$\varepsilon = \Phi\left(\frac{1+p}{2}\right) \cdot \sigma + \mu - \mu$$

Dies kann nun endgültig auf diesen Ausdruck gekürzt werden:

$$\varepsilon = \Phi^{-1}\left(\frac{1+p}{2}\right) \cdot \sigma$$

Allgemein gilt für die Bestimmung von x also:

p **für** $P(Z < X)$: $X = \Phi^{-1}(p) \cdot \sigma + \mu$

p **für** $P(Z > X)$: $X = -\Phi^{-1}(p) \cdot \sigma + \mu = \Phi^{-1}(1-p) \cdot \sigma + \mu$

Symmetrisches Intervall: $\varepsilon = \Phi^{-1}\left(\frac{1+p}{2}\right) \cdot \sigma = D^{-1}(p) \cdot \sigma$

μ bestimmen

Oftmals ist nur die Standardabweichung σ gegeben, ein Wert x der Zufallsvariable X sowie die Wahrscheinlichkeit dass p dass X größer oder kleiner als dieses x ist. Die Aufgabe besteht dann darin, μ zu bestimmen. Wiedermals wird diese Aufgabe durch eine Umformung der z-Transformationsformel gelöst:

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

$$\Downarrow$$

$$\mu = -z \cdot \sigma + x = -\Phi^{-1}(p) \cdot \sigma + x$$

Beispiel: *Bestimme den Erwartungswert μ so, dass bei einer Standardabweichung $\sigma = 2$ die normalverteilte Zufallsvariable X zu 95% einen Wert höher als 495 annimmt.*

Bei der Berechnung von μ muss man hier darauf achten, dass $p = P(Z > X)$. Das heißt, dass der z Wert entweder für die Gegenwahrscheinlichkeit $1 - p$ berechnet werden oder sein Vorzeichen verändert werden muss:

$$P(Z > X) \Rightarrow \Phi^{-1}(1-p) = -\Phi^{-1}(p)$$

μ wird dann so bestimmt:

$$\mu = -z \cdot \sigma + x = -[-\Phi^{-1}(p)] \cdot \sigma + x = 1.645 \cdot 2 + 495 = 498.3$$

Allgemein:

Wenn p **für** $P(X < x)$**:** $\mu = -z \cdot \sigma + x = -\Phi^{-1}(p) \cdot \sigma + x$

Wenn p **für** $P(X > x)$**:** $\mu = -z \cdot \sigma + x = -[-\Phi^{-1}(p)] \cdot \sigma + x = \Phi^{-1}(p) \cdot \sigma + x$

σ bestimmen

Zur Berechnung der Standardabweichung σ geht man ähnlich vor wie bei jener des Erwartungswertes:

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

$$\Downarrow$$

$$\sigma = \frac{x - \mu}{z} = \frac{x - \mu}{\Phi^{-1}(p)}$$

Beispiel: *Ermittle die Standardabweichung σ so, dass es ein symmetrisches Intervall $]\mu - \varepsilon; \mu + \varepsilon[$ mit $\varepsilon = 1$ geben kann, in welchem sich die Zufallsvariable X zu 95% befindet.*

Hierzu sollte klargestellt werden, dass der Ausdruck $x - \mu$ bei der z-Transformation den Abstand zwischen x und μ berechnet. Wir wissen hier, dass dieser symmetrische Abstand ε den Wert 1 hat, können $x - \mu$ also durch 1 ersetzen. Da es sich um ein symmetrisches Intervall handelt, können wir nicht direkt p verwenden, sondern müssen wie bei der Berechnung symmetrischer Intervalle die Wahrscheinlichkeit, dass Z geringer als die untere Intervallgrenze ist, zu p hinzufügen. Generell hilft uns hier die vorhin definierte Formel zur Berechnung eines symmetrischen Intervalls:

$$\varepsilon = \Phi^{-1}\left(\frac{1 + p}{2}\right) \cdot \sigma$$

$$\Downarrow$$

$$\sigma = \frac{\varepsilon}{\Phi^{-1}\left(\frac{1+p}{2}\right)} = \frac{1}{\Phi^{-1}(0.975)} = \frac{1}{1.955}$$

$$\Downarrow$$

$$\sigma = 0.51$$

Allgemein:

Wenn p für $P(Z < X)$: $\sigma = \frac{x - \mu}{\Phi^{-1}(p)}$

Wenn p für $P(Z > X)$: $\sigma = \frac{x - \mu}{\Phi^{-1}(1 - p)} = \frac{x - \mu}{-\Phi^{-1}(p)}$

Symmetrisches Intervall: $\sigma = \frac{\varepsilon}{\Phi^{-1}\left(\frac{1+p}{2}\right)} = \frac{\varepsilon}{D^{-}1(z)}$

Konfidenzintervalle

Ein Konfidenzintervall ist ein symmetrisches Intervall um einen Erwartungswert, in welchem die Werte der Zufallsvariable mit einer gewissen Wahrscheinlichkeit γ liegen. Diese Wahrscheinlichkeit wird *Sicherheit* γ genannt. Gegensätzlich zur normalen Berechnung symmetrischer Intervalle um den Erwartungswert, werden die Intervallgrenzen hier relativ und nicht absolut angegeben. Sie sind also Prozentsätze einer Grundmenge. Das bedeutet lediglich, dass die absoluten Werte, die das Resultat der normalen Berechnung sind, durch die Grundmenge dividiert werden müssen. Der Sinn dahinter ist es, anhand einer kleinen statistischen Stichprobe Schlüsse für eine größere Menge ziehen zu können.

Liegt die Zufallsvariable nämlich mit einer gewissen Sicherheit γ in einem relativen Intervall einer kleinen Menge, so wird sie dies höchst wahrscheinlich auch für eine grooßere Menge. Der Grund, wieso solche Schätzungen möglich sind, liegt darin, dass die Gauss'sche Glockenkurve bzw. die Normalverteilung für kleine Mengen eine ähnliche Form hat wie jene für große Mengen. Spricht man also nur von relativen Werten, so werden diese bei Stichproben *ähnlich* denen der Grundmenge sein. Dies ist beispielsweise bei Hochrechnungen von Wahlergebnissen günstig.

Beispiel: *Von 1500 Personen geben 390 Personen an, die Partei P zu wählen. Die Partei P möchte nun ein 99%-iges Konfidenzintervall für ihren Wähleranteil bei der nächsten Wahl wissen.*

Die Grundmenge n umfasst 1500 Personen. Von diesen geben 390 Personen an die Partei P zu wählen, die absolute Häufigtkeit h_A der Wähler der Partei P beträgt also 390. Die relative Häufigkeit h_R ist die absolute Häufigkeit, dividiert durch die Grundmenge. Dieser Wert ist gleichzeitig die Wahrscheinlichkeit p, mit der eine Person die Partei P wählt:

$$p = \frac{h_A}{n} = \frac{390}{1500} = 0.26$$

Somit ist die Anzahl der Wähler *binomialverteilt* mit $n = 1500$ und $p = 390$. Weil die Standardabweichung dieser binomialverteilten Zufallsvariable $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ mit ~ 17 größer 3 ist, darf sie durch eine Normalverteilung approximiert werden. Der Erwartungswert μ ist dabei gleich $n \cdot p$ und somit 390. Nun kann man ein normalverteiltes Modell $\mathcal{N}(\mu = 390; \sigma = 17)$ aufstellen. Gefragt ist jetzt, innerhalb welcher Intervallsgrenzen die normalverteilte Zufallsvariable X in 99% aller Fälle liegt. Da die Funktion $D(z)$ die Wahrscheinlichkeit p für einen symmetrisch um den Erwartungswert schwankenden z-Wert angibt, gibt ihr Invers $D^{-1}(p)$ den jeweiligen z-Wert für eine Wahrscheinlichkeit p an. Die Wahrscheinlichkeit des symmetrischen Intervalls ist hierbei $\gamma = 0.99$, weil der Wähleranteil zu γ Prozent in diesem Intervall liegen soll. Wüsste man die obere Intervallsgrenze x, würde man die Wahrscheinlichkeit γ, dass die Zufallsvariable in diesem Intervall liegt, so berechnen:

$$\gamma = D(z) = D(\frac{x-\mu}{\sigma})$$

Da man hier aber nur p kennt, formt man um:

$$x = D^{-1}(\gamma) \cdot \sigma + \mu$$

Somit erhält man die obere Grenze des Intervalls. Es ist aber der Abstand von μ gefragt, also kann man die Addition von μ weglassen:

$$x = D^{-1}(\gamma) \cdot \sigma$$

Setzt man ein:

$$x = D^{-1}(0.99) \cdot 17 = 2.576 \cdot 17 \approx 43.8$$

Die obere Grenze liegt nun also bei $\mu + x = 390 + 43.8 = 433.8$, die untere bei $\mu - x = 390 - 43.8 = 346.2$. Letztlich gilt es noch, die relativen Häufigkeiten zu bestimmen. Dafür dividiert man durch die Gesamtmenge $n = 1500$:

$$\frac{346.2}{1500} = 0.2308 \qquad \frac{433.8}{1500} = 0.2892$$

Das endgültige Konfidenzintervall mit der Sicherheit $\gamma = 0.99$ lautet dann:

$$[0.2308; 0.2892] = [23.08\%; 28.92\%]$$

Allgemein für den relativen, symmetrischen Abstand ε vom Erwartungswert μ, in welchem eine normalverteilte Zufallsvariable mit einer Sicherheit von γ Prozent liegt:

$$\varepsilon = \frac{D^{-1}(\gamma) \cdot \sigma}{n}$$