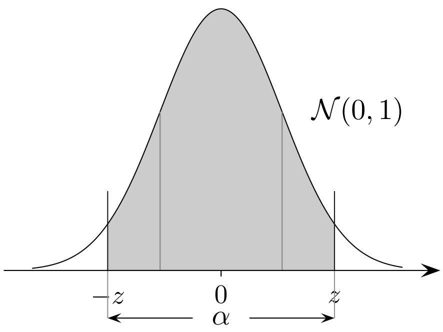
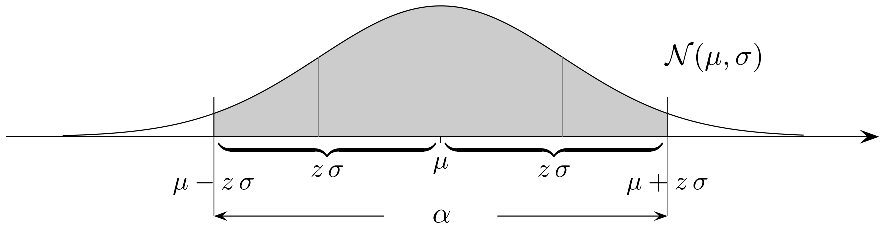
# Normalverteilung und Standardisierung




Die Normalverteilungen $\mathcal{N}(\mu, \sigma)$ ergeben sich aus der Standardnormalverteilung $\mathcal{N}(0,1)$ (Gaussche Glockenkurve) durch strecken und stauchen um $\sigma$ und verschieben um $\mu$.
Das Intervall $[-z, z]$ geht dabei in die $z\sigma$-Umgebung $[\mu - z\sigma,\ \mu + z\sigma]$ von $\mu$ über.
Einem $z$-Wert von $\mathcal{N}(0,1)$ entspricht $k = \mu + z\sigma$ von $\mathcal{N}(\mu, \sigma)$.
Umgekehrt wird dann einem $k$-Wert von $\mathcal{N}(\mu, \sigma)$ der Wert $z = \dfrac{k-\mu}{\sigma}$ zugeordnet.

Dies führt zu

$$P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

bei Approximation der Binomialverteilung wäre $\approx$ korrekter.

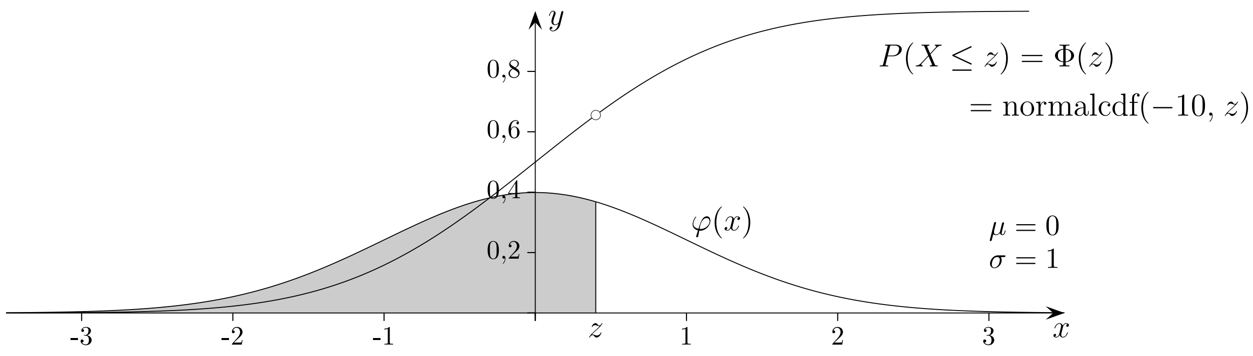
$$\underbrace{P(X \le k)}_{\beta} = \Phi\left(\frac{k-\mu}{\sigma}\right)$$

Falls $\beta$ gegeben ist, folgt

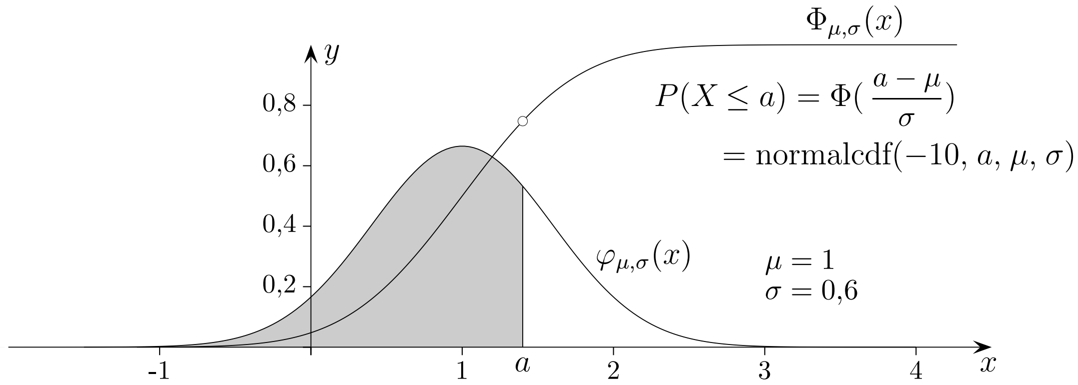
$$\frac{k-\mu}{\sigma} = \Phi^{-1}(\beta)$$

Dieses kann je nach Fragestellung nach $k$, $\mu$ oder $\sigma$ umgestellt werden.

GTR
$\Phi^{-1}(\beta) = \text{invNorm}(\beta)$




$\Phi(x) = \int\limits_{-\infty}^{x} \varphi(t)\,dt$ ist eine monoton wachsende Integralfunktion.
Somit existiert die Umkehrfunktion $\Phi^{-1}(y) = \text{invNorm}(y)$ für die Standardnormalverteilung, bzw. $\text{invNorm}(y, \mu, \sigma)$ für beliebige Normalverteilungen, um zu gegebener Wahrscheinlichkeit $y$, $0 < y < 1$, den $z$-Wert zu ermitteln.

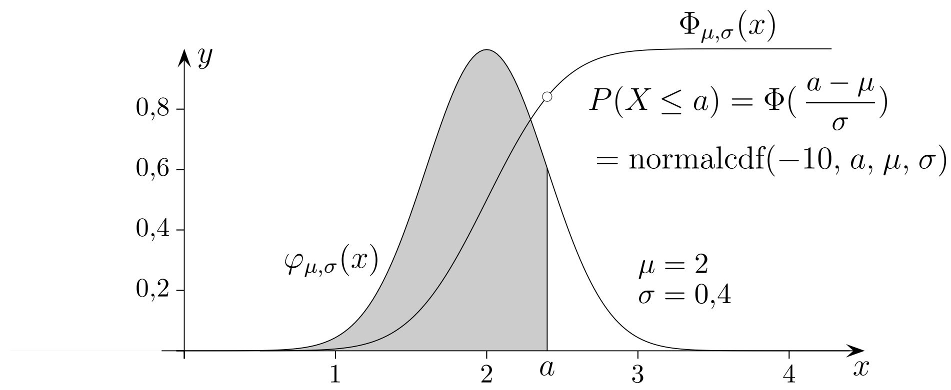



Aus $a = \mu + z\sigma$ folgt

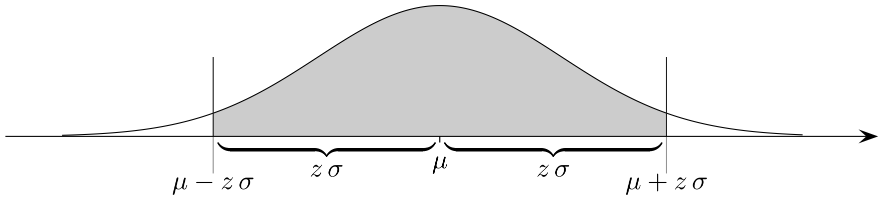
$$z = \frac{a-\mu}{\sigma}.$$

Der Inhalt der Fläche unter $\varphi_{\mu,\sigma}(x)$
bis zur rechten Grenze $a$ ist daher

$$\Phi_{\mu,\sigma}(a) = \Phi(\frac{a-\mu}{\sigma}).$$



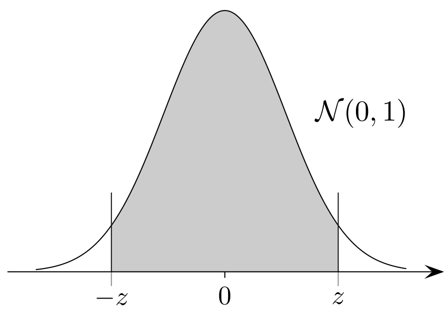
# $z\sigma$-Umgebung

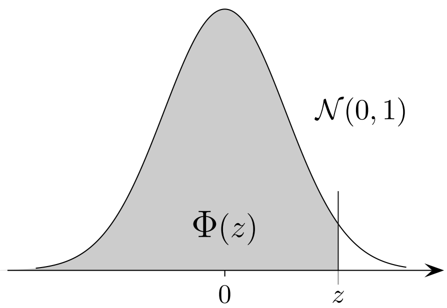



$$P(a \leq X \leq b) \;=\; \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(\mu - z\,\sigma \leq X \leq \mu + z\,\sigma) \;=\; \Phi(z) \;-\; \Phi(-z) \;=\; 2\,\Phi(z) - 1$$

Erläutere die letzte Zeile.



a) Sei $P(X \leq 15) = 20\,\%$, $\mu = 25$
gesucht $\sigma$

b) Sei $P(X \leq 20) = 10\,\%$, $\sigma = 5$
gesucht $\mu$

c) $P(X \leq 30) = 70\,\%$
$P(X \leq 24) = 10\,\%$
$\mu$, $\sigma = ?$

a) Sei $P(X \leq 15) = 20\,\%,\ \mu = 25$
gesucht $\sigma$

Lösung

$$\Phi\left(\frac{15-\mu}{\sigma}\right) = 0{,}2$$

$$\frac{15-\mu}{\sigma} = \Phi^{-1}\big(0{,}2\big)$$

$$\sigma = \frac{15-\mu}{\Phi^{-1}\big(0{,}2\big)} = 11{,}9$$

b) Sei $P(X \leq 20) = 10\,\%,\ \sigma = 5$
gesucht $\mu$

Lösung

$$\Phi\left(\frac{20-\mu}{\sigma}\right) = 0{,}1$$

$$\frac{20-\mu}{\sigma} = \Phi^{-1}\big(0{,}1\big)$$

$$\mu = 20 - \sigma \cdot \Phi^{-1}\big(0{,}1\big) = 26{,}4$$

c) $P(X \leq 30) = 70\,\%$
$P(X \leq 24) = 10\,\%$
$\mu,\ \sigma = ?$

Lösung

$$P(Y \leq 30) \ = \ \Phi\left(\frac{30-\mu}{\sigma}\right) = 70\,\%$$

$$P(Y \leq 24) \ = \ \Phi\left(\frac{24-\mu}{\sigma}\right) = 10\,\%$$

$$\Phi^{-1}\big(0{,}70\big) = 0{,}524$$

$$\Phi^{-1}\big(0{,}10\big) = -1{,}282$$

$\mu = 28{,}3$
$\sigma = 3{,}3$

# Normalverteilung, GTR

a) $P(X \leq b) = 80\,\%,\ \mu = 50,\ \sigma = 3$ (in $g$)
gesucht $b$

b) Sei $P(1 \leq X \leq 2) = 20\,\%,\ \sigma = 0{,}5$
gesucht $\mu$

c) Sei $P(X \leq 20) = 15\,\%,\ \sigma = 5$ (in $g$)
gesucht $\mu$

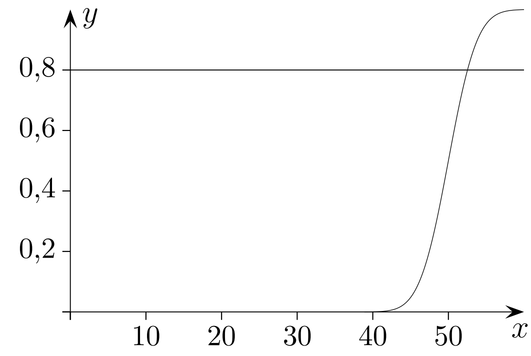
d) Sei $P(X \leq 40) = 80\,\%,\ \mu = 36$ (in $g$)
gesucht $\sigma$

e) $P(X \leq 34) = 12\,\%$ ($X$ in $g$)
$P(X \geq 44) = 5\,\%$
$\mu,\ \sigma = ?$

Wenn die Einheit z. B. Gramm ist, wird $X \geq 0$ vorausgesetzt.
Die Verwendung der Normalverteilung impliziert dann $P(0 \leq X \leq \mu) = 0{,}5$.
Die linke Grenze für $X$ in $P(X \leq a)$ ist in diesem Fall also 0.

# Normalverteilung, GTR

a) $P(X \leq b) = 80\,\%,\ \mu = 50,\ \sigma = 3$ (in $g$)
gesucht $b$

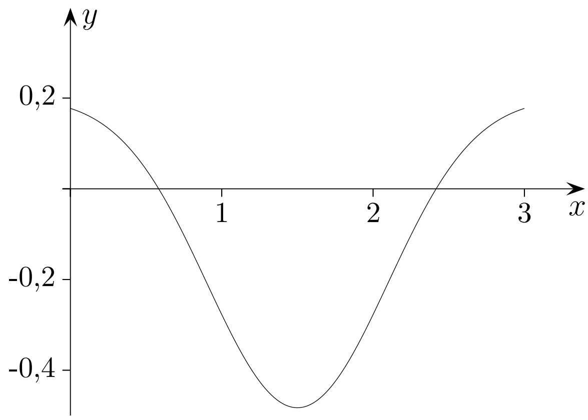



Schnittstelle von
$\backslash\mathrm{Y}_1 = \mathrm{normalcdf}(0, \mathrm{X}, 50, 3)$
$\backslash\mathrm{Y}_2 = 0{,}8$

$b = 52{,}52$

Jedoch geht es einfacher mit $b = \mathrm{invNorm}(0.8, 50, 3)$.

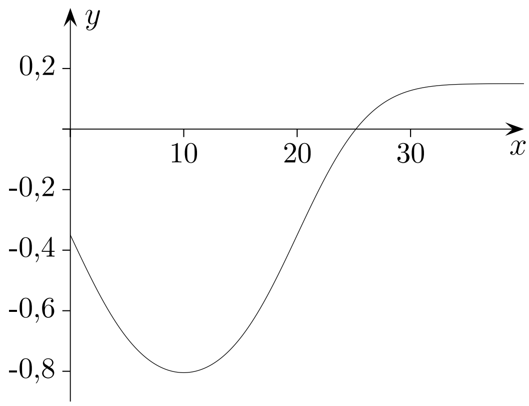
b) Sei $P(1 \leq X \leq 2) = 20\,\%,\ \sigma = 0{,}5$
gesucht $\mu$




Nullstellen von
$\backslash\mathrm{Y}_1 = 0.2 - \mathrm{normalcdf}(1{,}2{,}\mathrm{X}{,}0.5)$

$\mu_1 = 0{,}583$
$\mu_2 = 2{,}417$

c) Sei $P(X \leq 20) = 15\,\%,\ \sigma = 5$ (in $g$)
gesucht $\mu$

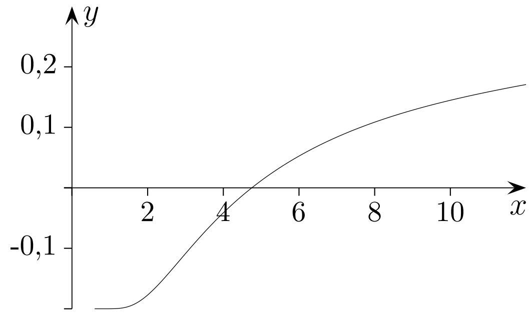



Nullstelle von
$\backslash\mathrm{Y}_1 = 0.15 - \mathrm{normalcdf}(0{,}20{,}\mathrm{X}{,}5)$

$\mu = 25{,}182$

# Normalverteilung, GTR

d) Sei $P(X \leq 40) = 80\,\%$, $\mu = 36$ (in $g$)
gesucht $\sigma$



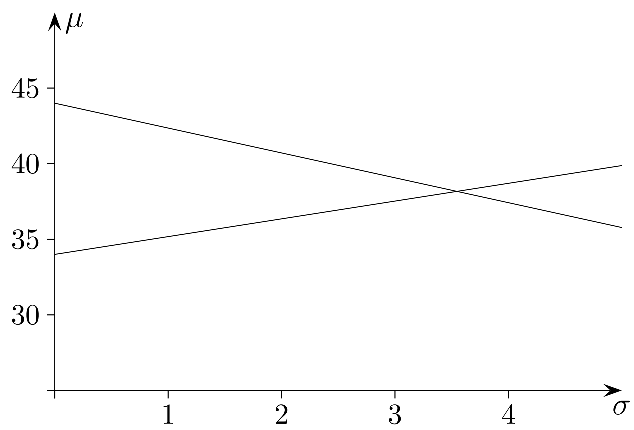

Nullstelle von
$\backslash\mathrm{Y}_1 = 0.80 - \mathrm{normalcdf}(0,40,36,\mathrm{X})$
$\sigma = 4{,}753$

e) $P(X \leq 34) = 12\,\%$ ($X$ in $g$)
$P(X \geq 44) = 5\,\%$
$\mu$, $\sigma = ?$

$$P(Y \leq 34) \;=\; \Phi\left(\frac{34-\mu}{\sigma}\right) = 12\,\% \quad\Longrightarrow\quad \mu = 34 - \Phi^{-1}\big(0{,}12\big)\cdot\sigma$$

$$P(Y \geq 44) \;=\; 1 - \Phi\left(\frac{44-\mu}{\sigma}\right) = 5\,\% \quad\Longrightarrow\quad \mu = 44 - \Phi^{-1}\big(0{,}95\big)\cdot\sigma$$

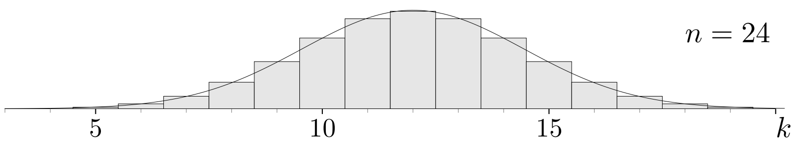
Schnittpunkt
$\mu = 38{,}167$
$\sigma = 3{,}546$

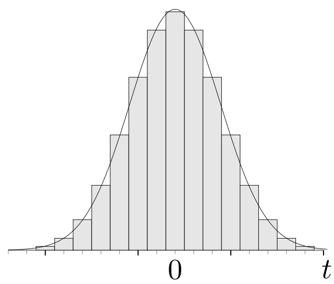



Nur b) muss grafisch-numerisch bearbeitet werden.
Für a), c) und d) führt die Verwendung von $\Phi^{-1}$ zur Lösung,
in e) auch ein LGS.

# Standardisierte Normalverteilung

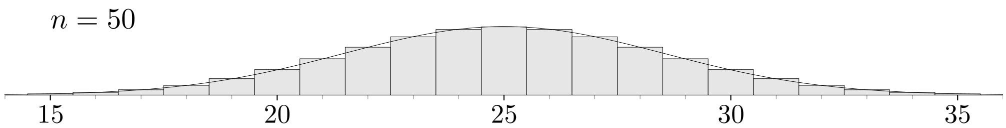
Um das Verhalten von Binomialverteilungen für $n \to \infty$ zu untersuchen (hier für $p = \frac{1}{2}$ dargestellt), werden die $k$-Werte in ihrer relativen Lage zum Erwartungswert $\mu$ betrachtet.

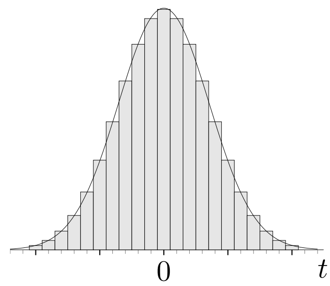


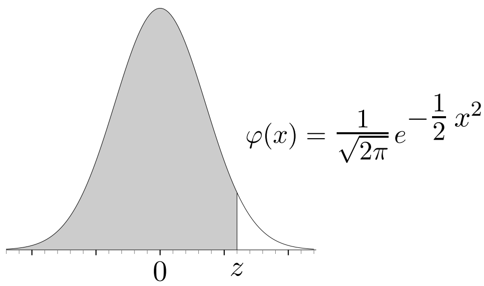

$$
\begin{aligned}
k &= \mu + t \cdot \sigma \\
\iff \quad k - \mu &= t \cdot \sigma \\
\iff \quad t &= \frac{k-\mu}{\sigma}
\end{aligned}
$$

Der $t$-Wert gibt an, um welches Vielfache von $\sigma$ $k$ vom Erwartungswert $\mu$ abweicht.
Für $n \to \infty$ braucht man nur kleine $t$-Werte zu berücksichtigen.
Um über der $t$-Achse ein Histogramm aufzutragen, werden die Rechtecksbreiten durch $\sigma$ dividiert.
Damit die Flächeninhalte gleich bleiben, müssen die Höhen mit $\sigma$ multipliziert werden.




Nun ist absehbar, was sich für $n \to \infty$ ergibt.
Die standardisierten Verteilungen werden durch die Gausssche Glockenkurve $\varphi(x)$ approximiert.
Deren Verteilungsfunktion $\Phi(z)$ gibt den Flächeninhalt unter der Kurve von $-\infty$ bis $z$ an.

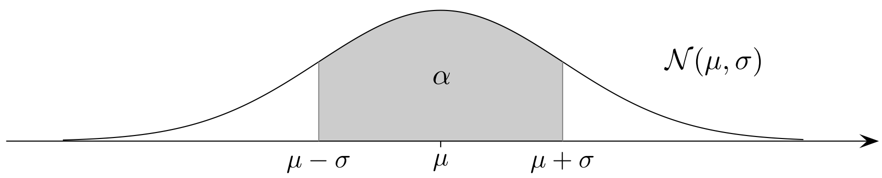



Wir erhalten:

$$P(a \le X \le b) \approx \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

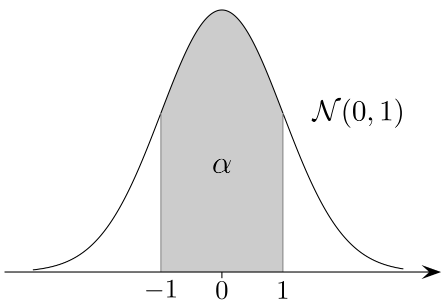
(ohne Stetigkeitskorrektur)

# $\sigma$-Umgebung




Wie groß ist $\alpha$?

Antwort:
$\mathcal{N}(\mu, \sigma)$ geht aus $\mathcal{N}(0, 1)$ durch Verschieben, Strecken und Stauchen hervor.
0 geht in $\mu$ über, $-1$ in $\mu - \sigma$, 1 in $\mu + \sigma$.




$\alpha = \text{normalcdf}(-1, 1) = 68{,}3\,\%$

Weiter erhalten wir:

$P(2\sigma\text{-Umgebung}) = \text{normalcdf}(-2, 2) = 95{,}4\,\%$

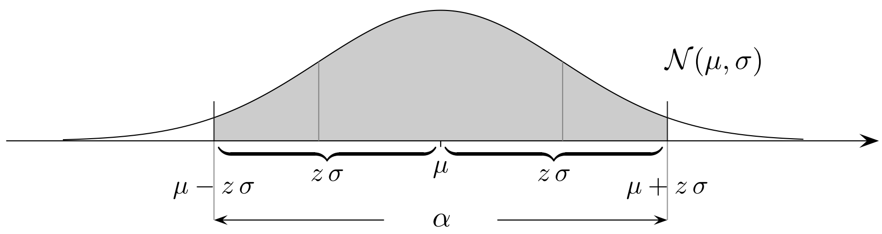
$P(3\sigma\text{-Umgebung}) = \text{normalcdf}(-3, 3) = 99{,}7\,\%$

Sei für eine Binomialverteilung $n = 5000$, $p = 0{,}5$ gegeben.
Welche Trefferzahlen sind mit 95,4 %iger Wahrscheinlichkeit zu erwarten?
(Der Bereich heißt auch Schwankungsintervall.)

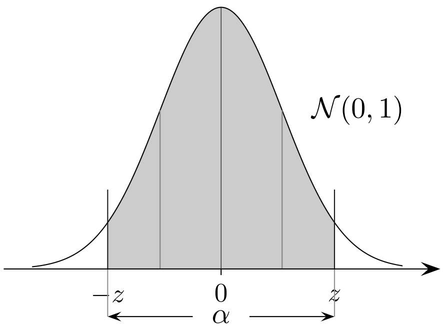
Sei für eine Binomialverteilung $n = 5000$, $p = 0{,}5$ gegeben.
Welche Trefferzahlen sind mit 95,4 %iger Wahrscheinlichkeit zu erwarten?
(Der Bereich heißt auch Schwankungsintervall.)

$[2430; 2570]$

# $z\sigma$-Umgebung



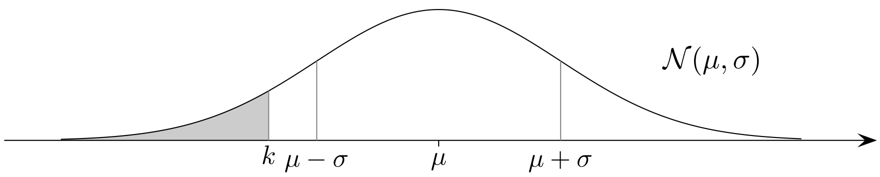

$\alpha = 90\,\%$ (z. B.) gegeben,
$z$ gesucht



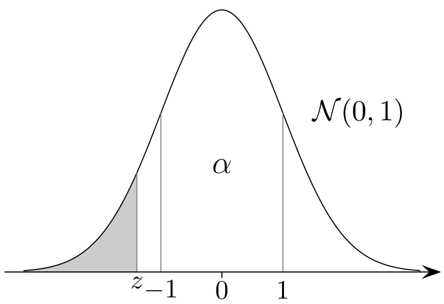

Begründe:

$$z = \underbrace{\text{invNorm}}_{\Phi^{-1}}\left(\frac{1}{2} + \frac{\alpha}{2}\right) = \text{invNorm}\left(\frac{1+\alpha}{2}\right)$$

# $\mu$, $\sigma$ ermitteln




$P(X \leq 235) = 9\,\%,\ \ \mu = 250$
gesucht $\sigma$




Welcher Zusammenhang besteht zwischen $k$ und $z$?

Antwort:

$$\frac{k-\mu}{\sigma} = z \qquad \text{(Setze für } k\ \ \mu - \sigma\text{, bzw. } \mu + \sigma \text{ ein.)}$$

$$\underbrace{\text{normalcdf}}_{\Phi}\left(\frac{k-\mu}{\sigma}\right) = P(X \leq k) \qquad \mid \Phi^{-1}$$

...

$\sigma = 11{,}2$