Lösung zu Aufgabe 3: Normalverteilte Messgrößen

a) Anteil der Stifte mit $d_i \leq 4{,}3$ mm:

Die Wahrscheinlichkeit $P(x_1 < x \leq x_2)$ *dafür, dass ein Messwert* x *im Intervall* $x_1 < x < x_2$ *liegt, kann durch das Integral*

$$P(x_1 < x \leq x_2) = \int_{x_1}^{x_2} h(x)dx = \lim_{n\to\infty} \frac{\Delta n}{n}\Bigg|_{x_1}^{x_2}$$

berechnet werden. Dies ist die Fläche unter der Kurve $h(x)$ *im Intervall* $(x_1 < x < x_2)$.

Die Fläche unter der gesamten Kurve ist wie die Fläche unter dem Histogramm gleich 1, d.h.:

$$P(-\infty < x \leq \infty) = 1$$

Die Wahrscheinlichkeitsfunktion oder Verteilungsfunktion $P(x)$ *gibt die Wahrscheinlichkeit dafür an, dass ein Messwert* x_a *kleiner oder gleich einer Schranke* x *ist.*

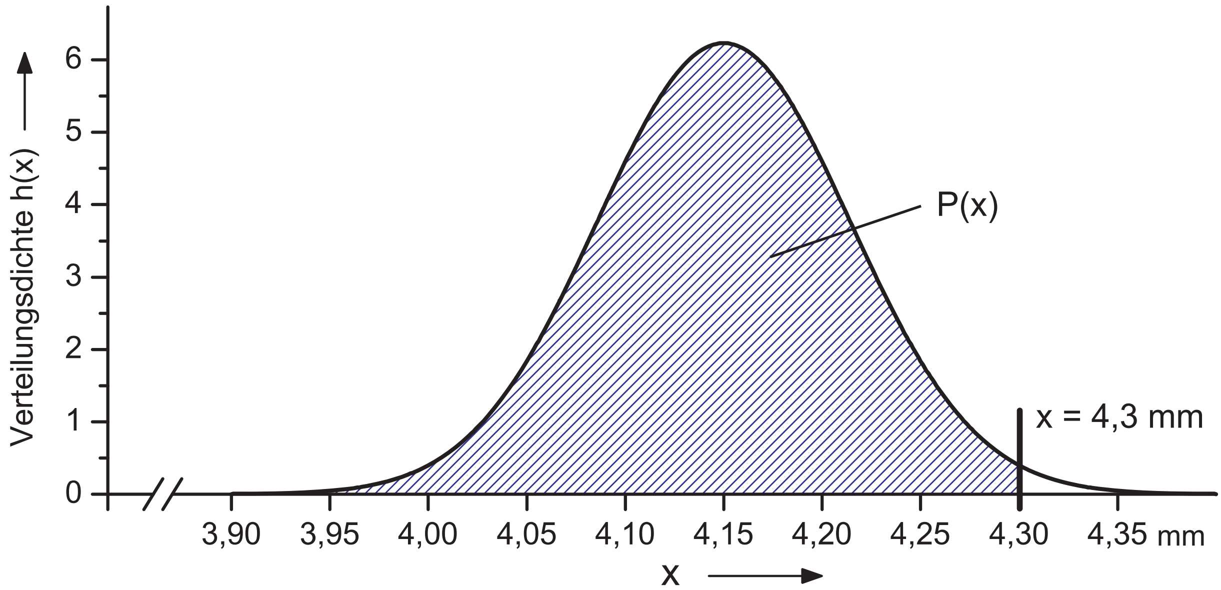
$$P(x) = P(x_a \leq x) = \int_{-\infty}^{x} h(x_a)dx_a$$

Im vorliegenden Fall handelt es sich bei der Verteilungsdichtefunktion um eine Gaußsche Normalverteilung mit einem Erwartungswert von $\mu = 4{,}15$ mm und einer Standardabweichung von $\sigma = 0{,}064$ mm.

In Aufgabeteil a) ist nun die Wahrscheinlichkeit für den Fall gesucht, dass ein Messwert x_a kleiner oder gleich der Schranke x = 4,3 mm ist. Entsprechend obiger Notation suchen wir als die Wahrscheinlichkeit P(x) für den Fall x = 4,3 mm:

$$P(4{,}3\text{ mm}) = P(x_a \leq 4{,}3\text{ mm}) = ?$$

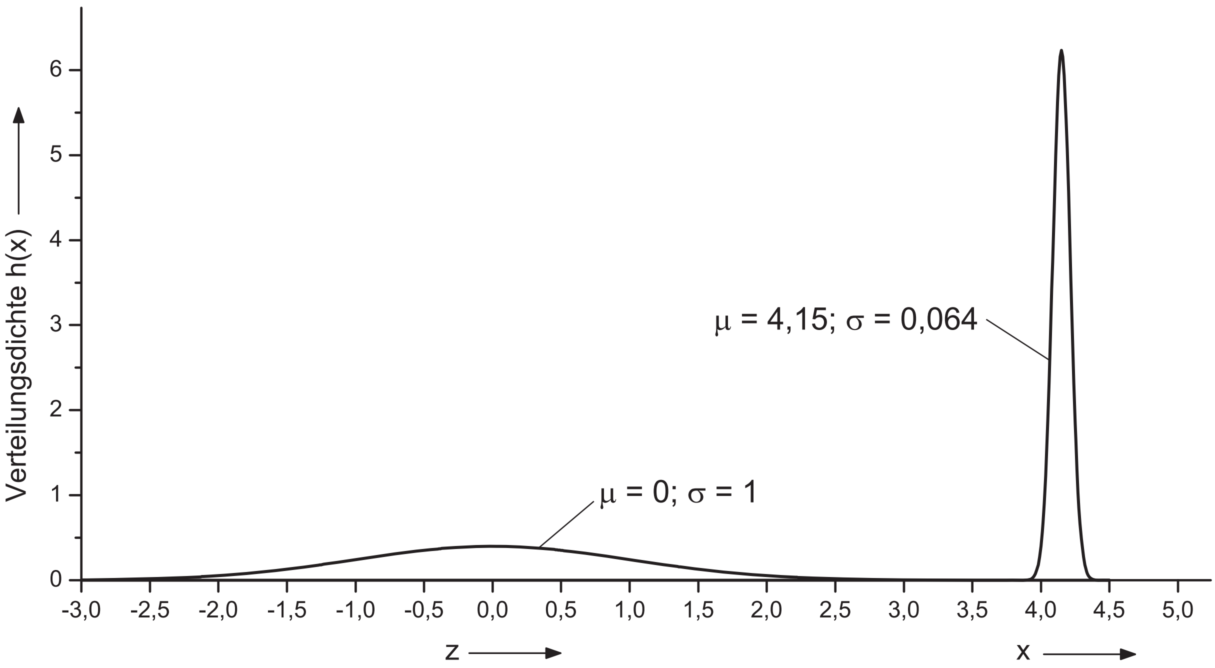
Die nachfolgende Abbildung veranschaulicht die Problemstellung. Die schraffierte Fläche kennzeichnet darin das Integral unter der Verteilungsdichtefunktion der gegebenen Gaußschen Normalverteilung in den Grenzen von $-\infty$ bis x = 4,3 mm.



Die Wahrscheinlichkeitsfunktion $P(x)$ *der Gaußschen Normalverteilung ist nicht als geschlossene Funktion darstellbar. Sie ist als grafische Darstellung bzw. tabellarisch in statistischen Handbüchern zu finden.*

Entsprechende Tabellen liegen naturgemäß jedoch nicht für die jeweilige Normalverteilung mit speziellen Werten von Erwartungswert und Standardabweichung vor. Stattdessen geben diese Tabellen die Wahrscheinlichkeitsfunktion $P(x)$ in normierten Koordinaten an. Die entsprechend normierte Gaußsche Normalverteilung weist die Parameter $\mu = 0$ und $\sigma = 1$ auf und wird als standardisierte Normalverteilung bezeichnet.

Um anhand einer derartigen Tabelle Aussagen über die in dieser Aufgabe betrachtete Verteilung treffen zu können, muss also eine entsprechende Normierung vorgenommen werden. Anschaulich entspricht dies der Anwendung einer Berechnungsvorschrift, welche die gegebene Normalverteilung der Größe x in eine standardisierte Normalverteilung der Größe z transformiert. In nachfolgender Abbildung sind die beiden Verteilungen – die standardisierte Normalverteilung mit $\mu = 0$ und $\sigma = 1$ sowie die betrachtete spezielle Normalverteilung mit $\mu = 4{,}15$ und $\sigma = 0{,}064$ – in einem gemeinsamen Diagramm aufgetragen.



Um nun die spezielle Normalverteilung in eine standardisierte Normalverteilung zu transformieren, muss erstens eine Verschiebung um $-\mu$ und zweitens eine Spreizung um den Faktor $\frac{1}{\sigma}$ vorgenommen werden. Die transformierte Koordinate z der standardisierten Normalverteilung ergibt sich aus der Koordinate x der speziellen Normalverteilung also gemäß folgendem Zusammenhang:

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

Im vorliegenden Fall wird also zu dem, der x-Koordinate entsprechenden, oberen Durchmesser d_i = 4,15 mm die entsprechende, auf die standardisierte Normalverteilung bezogenen z-Koordinate gesucht. Mit den bekannten Parametern $\mu = 4{,}15$ und $\sigma = 0{,}064$ der vorliegenden Verteilung ergibt sich daher:

$$z = \frac{4{,}3 - 4{,}15}{0{,}064} = 2{,}34375$$

Da in der uns zur Verfügung stehenden Tabelle der Summenfunktion der standardisierten Normalverteilung (siehe Übungsskript S. 19) die Schrittweite der eingetragenen z-Werte $\Delta z = 0{,}01$ beträgt, ist es sinnvoll, die berechneten z-Werte auf zwei Nachkommastellen zu runden. Es folgt somit:

$$z = 2{,}34$$

Zu diesem z-Wert gilt es nun, aus der besagten Tabelle die relative Summenhäufigkeit $\Phi(z)$ zu ermitteln. Die von uns genutzte Tabelle weist in Zeilenrichtung eine Schrittweite von $\Delta z = 0{,}1$ und in Spaltenrichtung eine Schrittweite von $\Delta z = 0{,}01$ auf. Der gesuchte Wert $\Phi(z)$ befindet sich an der Position innerhalb der Tabelle, für welche die Summe aus dem am

Zeilenanfang und dem am Spaltenanfang verzeichneten z-Werten dem vorliegenden z-Wert entspricht (vergleiche hierzu auch das Ablesebeispiel im Kopf der Tabelle).

Im vorliegenden Fall finden wir den Wert $\Phi(z = 2{,}34)$ am Schnittpunkt der Zeile mit dem Wert 2,3 und der Spalte mit dem Wert 0,04. Der dort verzeichnete Wert der Summenfunktion lautet:

$$\Phi(z = 2{,}34) = 0{,}990358$$

Dieser Zahlenwert gibt wie eingangs erläutert eine Wahrscheinlichkeit an. Als prozentuale Wahrscheinlichkeit ausgedrückt, entspricht dieses Ergebnis daher:

$$\Phi(z = 2{,}34) = 99{,}0358\%$$

Die Antwort auf die eingangs gestellte Frage lautet somit:

Bei den vorgegebenen Prozessparametern weisen rund 99,04% aller gefertigten Stifte einen Durchmesser von $d_i \leq 4{,}3$ mm auf!

b) Wahrscheinlichkeit für einen Stift mit 4,09 mm ≤ d_i ≤ 4,23 mm:

In Aufgabeteil b) ist nun die Wahrscheinlichkeit für den Fall gesucht, dass ein Messwert x_a größer oder gleich einer unteren Schranke von x_u = 4,09 mm und kleiner oder gleich einer oberen Schranke von x_o = 4,23 mm ist:

$$P(x_u \leq x_a \leq x_o) = ?$$

oder mit den vorliegenden Zahlenwerten:

$$P(4{,}09\,\text{mm} \leq x_a \leq 4{,}23\,\text{mm}) = ?$$

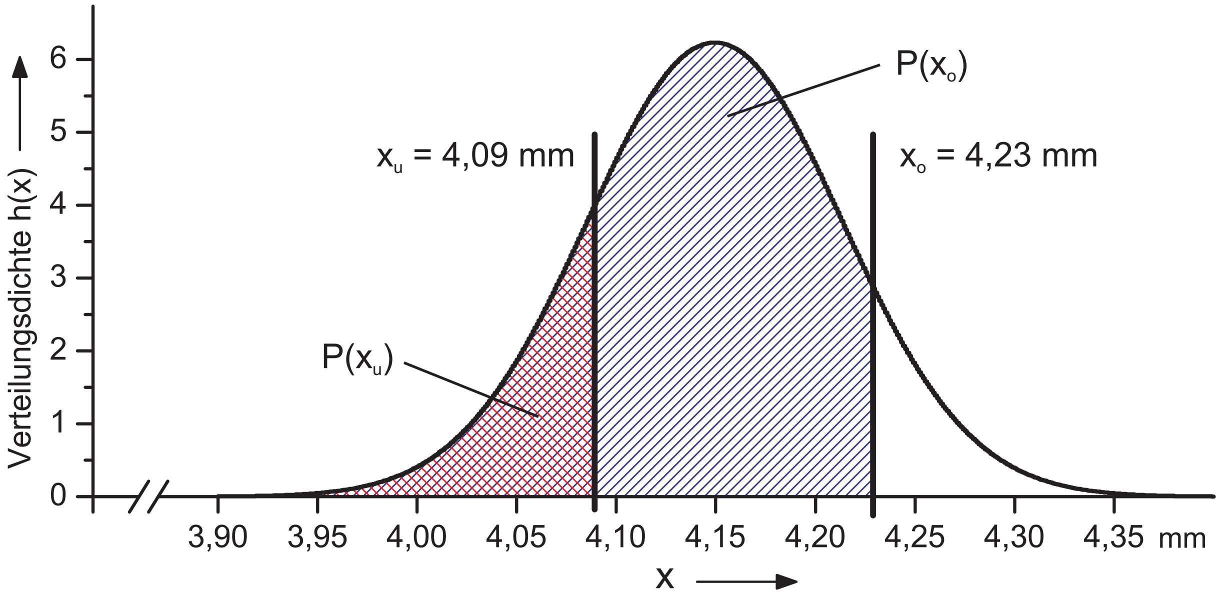
Wie anschaulich nachvollziehbar ist, lässt sich die gesuchte Wahrscheinlichkeit dadurch ermitteln, dass zunächst separat die Wahrscheinlichkeiten für die beiden Fälle ermittelt werden, dass ein Messwert x_a kleiner oder gleich der oberen Schranke bzw. kleiner oder gleich der unteren Schranke ist und anschließend die Differenz dieser beiden Wahrscheinlichkeiten ermittelt wird. Wir können also allgemein schreiben:

$$P(x_u \leq x_a \leq x_o) = P(x_o) - P(x_u)$$

Mit den vorgegebenen Zahlenwerten ergibt sich also:

$$P(4{,}09\,\text{mm} \leq x_a \leq 4{,}23\,\text{mm}) = P(x_a \leq 4{,}23\,\text{mm}) - P(x_a \leq 4{,}09\,\text{mm})$$

Die nachfolgende Abbildung veranschaulicht diesen Lösungsansatz. Die blau schraffierte Fläche kennzeichnet darin das Integral unter der Verteilungsdichtefunktion der gegebenen Gaußschen Normalverteilung in den Grenzen von $-\infty$ bis x_o = 4,23 mm, die rot schraffierte Fläche entspricht dem Integral in den Grenzen von $-\infty$ bis x_u = 4,09 mm.



Die Berechnung der Wahrscheinlichkeiten $P(x_o)$ und $P(x_u)$ erfolgt analog zu Aufgabenteil a). Für die obere Schranke x_o = 4,23 mm ergibt sich daher:

$$z_o = \frac{x_o - \mu}{\sigma} \quad \Rightarrow \quad z_o = \frac{4,23 - 4,15}{0,064} = 1,25$$

Für die untere Schranke x_u = 4,09 mm gilt entsprechend:

$$z_u = \frac{x_u - \mu}{\sigma} \quad \Rightarrow \quad z_u = \frac{4,09 - 4,15}{0,064} = -0,9375 \approx -0,94$$

Aus der Tabelle der Summenfunktion der standardisierten Normalverteilung erhalten wir die zur oberen Schranke z_o gehörige Summenhäufigkeit:

$$\Phi(z_o = 1,25) = 0,89435$$

Zu der unteren Schranke finden wir in der Tabelle zunächst keinen passenden Wert, da die Tabelle nur die Summenhäufigkeiten für z-Werte im Bereich von 0 bis 2,99 auflistet. Da es sich bei der zugrundeliegenden Verteilungsdichtefunktion der standardisierten Normalverteilung jedoch um eine zur Koordinate Null symmetrische Funktion handelt, lassen sich durch Ausnutzung der Symmetrie daraus auch die Summenhäufigkeiten für z-Werte im Bereich von –2,99 bis 0 ableiten. Wie im Kopf der Tabelle aufgeführt, gilt die Symmetriebedingung:

$$\Phi(z) = 1 - \Phi(-z)$$

Für den vorliegenden Wert von $z_u = -0,94$ gilt also:

$$\Phi(-0,94) = 1 - \Phi(0,94)$$

Mit Hilfe der Tabelle ergibt sich somit:

$$\Phi(z_u = -0,94) = 1 - 0,826391 = 0,173609$$

Die gesuchte Wahrscheinlichkeit $P(4,09\text{ mm} \le x_a \le 4,23\text{ mm})$ ergibt sich damit zu:

$$\begin{aligned} P(4{,}09\,\text{mm} \le x_a \le 4{,}23\,\text{mm}) &= \Phi(z_o) - \Phi(z_u) \\ &= 0{,}89435 - 0{,}173609 \\ &= 0{,}720741 \\ &\mathrel{\hat{=}} 72{,}07\% \end{aligned}$$

Die Antwort auf die eingangs gestellte Frage lautet somit:

Die Wahrscheinlichkeit dafür, dass bei einer Stichprobe der Durchmesser eines einzeln entnommenen Stiftes im Bereich 4,09 mm ≤ d_i ≤ 4,23 mm liegt, beträgt rund 72,07%!

c) Anteil der Stifte im Bereich 4,086 mm ≤ d_i ≤ 4,214 mm:

Prinzipiell lässt sich Aufgabenteil c) rechnerisch analog zu Aufgabenteil b) lösen. Bei den gegebenen Werten bietet sich jedoch eine „elegantere" Lösung an, welche eine explizite Berechnung überflüssig macht.

Es fällt auf, dass das betrachtete Intervall [4,086 mm; 4,214 mm] symmetrisch zum Erwartungswert der vorgegebenen Verteilung von μ = 4,15 mm liegt, und dass ferner die Intervallgrenzen um den Betrag der Standardabweichung von σ = 0,064 mm gegenüber dem Erwartungswert verschoben sind.

$$[4{,}086\,\text{mm}; 4{,}214\,\text{mm}] = [4{,}15\,\text{mm} - 0{,}064\,\text{mm}; 4{,}15\,\text{mm} + 0{,}064\,\text{mm}]$$

Allgemein lässt sich das Intervall also wie folgt ausdrücken:

$$[4{,}086\,\text{mm}; 4{,}214\,\text{mm}] \mathrel{\hat{=}} [\mu - \sigma; \mu + \sigma]$$

Wie nun aus der Vorlesung bekannt ist (vgl. Vorlesungsskript, Tabelle 2.2), liegen bei einer normalverteilten Größe in einem Intervall von $\pm\sigma$ um den Erwartungswert μ stets 68,3% aller Werte.

Die Antwort auf die eingangs gestellte Frage lautet somit:

Bei den vorgegebenen Prozessparametern weisen 68,3% aller gefertigten Stifte einen Durchmesser im Bereich von 4,086 mm ≤ d_i ≤ 4,214 mm auf!

Anmerkung: Da in der Praxis meist Konfidenzintervalle angegeben werden, die eine Breite von $\pm\sigma$, $\pm 2\sigma$ oder $\pm 3\sigma$ symmetrisch zum Erwartungswert μ aufweisen, ist es lohnenswert, sich die zugehörigen Wahrscheinlichkeiten zu merken:

$$[\mu - \sigma; \mu + \sigma] \quad \rightarrow \quad 68{,}3\%$$

$$[\mu - 2\sigma; \mu + 2\sigma] \quad \rightarrow \quad 95{,}45\%$$

$$[\mu - 3\sigma; \mu + 3\sigma] \quad \rightarrow \quad 99{,}73\%$$

d) Benötigte Standardabweichung für mindestens 80% der Stifte im Bereich 4,086 mm ≤ d_i ≤ 4,214 mm:

Da, wie bereits unter Aufgabenteil c) festgestellt, das betrachtete Intervall symmetrisch zum Erwartungswert μ liegt, kann auch hier die Berechnung vereinfacht werden, wenn die dadurch bedingte Symmetrie ausgenutzt wird.

Wenn innerhalb des Intervalls 80% aller Messwerte liegen sollen, liegen also 20% außerhalb des Intervalls. Aufgrund der Symmetrie bedeutet dies zugleich, dass 10% der Werte kleiner als die untere Schranke von d_{min} = 4,086 mm und ebenfalls 10% größer als die obere Schranke von d_{max} = 4,214 mm sind.

Wir können im vorliegenden Fall daher die Betrachtung auf eine der beiden Grenzen des Intervalls beschränken. Da die vorliegende Tabelle positive z-Werte auflistet – also Grenzen, die größer als der Erwartungswert sind – wählen wir der Einfachheit halber die obere Grenze von d_{max} = 4,214 mm.

Aus unserer oben angestellten Betrachtung wissen wir nun, dass die Summenhäufigkeit dieser oberen Grenze $\Phi = 0,9$ beträgt, dass also im Intervall von $-\infty$ bis d_{max} 90% aller Werte liegen. Zunächst müssen wir den zu dieser Wahrscheinlichkeit gehörigen z-Wert ermitteln. Wir suchen also ein z, für das gilt:

$$\Phi(z) \overset{!}{=} 0,9$$

Eine Möglichkeit, diesen Wert zu bestimmen, besteht darin, die bereits zuvor verwendete Tabelle der Summenfunktion rückwärts abzulesen, also in der Tabelle den Φ-Wert zu suchen, welcher am nächsten an dem vorliegenden Wert von 0,9 liegt und aus der Zeilen- und Spaltenposition den zugehörigen z-Wert zu ermitteln.

Einfacher ist es im vorliegenden Fall jedoch, die kleine separate Tabelle am Fuß der Haupttabelle zu nutzen. Dort sind für einige glatte Werte von Φ(z) die zugehörigen z-Werte aufgeführt. Dort finden wird die Zuordnung:

$$\Phi(z) = 90\% \quad \rightarrow \quad z = 1,282$$

Um nun auf die gesuchte Standardabweichung zu kommen, nutzen wir die oben eingeführte Gleichung:

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

Da im vorliegenden Fall die Größen z, μ und x bekannt sind, formen wir nach der gesuchten Größe σ um:

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

Mit den gegebenen bzw. berechneten Größen

$x = d_{max} = 4{,}214$ mm

$\mu = 4{,}15$ mm

$z = 1{,}282$

ergibt sich für die gesuchte Standardabweichung σ somit:

$$\sigma = \frac{4{,}214\ \text{mm} - 4{,}15\ \text{mm}}{1{,}282} \approx 0{,}0499\ \text{mm}$$

Die Antwort auf die eingangs gestellte Frage lautet somit:

Damit bei gleichem Erwartungswert μ mindestens 80% der Positionierstifte im Intervall 4,086 mm ≤ d_i ≤ 4,214 mm liegen, müsste die Standardabweichung σ auf rund 0,0499 mm verbessert werden!

e) Betrachtung der Mittelwerte aus Stichproben vom Umfang n:

Führt man bei einer mit den Parametern μ und σ normalverteilten Messgröße nacheinander viele Messreihen vom Umfang n unter gleichen Randbedingungen durch, so kann man feststellen, dass deren Mittelwerte $\overline{x}_i$ eine um den Faktor $\frac{1}{\sqrt{n}}$ kleinere Streuung $S_{\overline{x}}$ aufweisen, als die Einzelmesswerte. Dieser Sachverhalt lässt sich – wie im Vorlesungsskript eingeführt – mit Hilfe der Abweichungsfortpflanzung für zufällige Abweichungen herleiten. Allgemein gilt für den Erwartungswert und die Standardabweichung des Mittelwertes:

$$\mu_{\overline{x}} = \mu$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Man beachte hierbei den Einfluss der Anzahl der Messungen auf die Standardabweichung. Misst man zum Beispiel viermal statt einmal so halbiert sich die Standardabweichung des Mittelwertes.

Im vorliegenden Fall, bei einer Messgröße x mit einer Standardabweichung von σ = 0,064 mm, ergibt sich für die Standardabweichung des Mittelwertes von Stichproben vom Umfang n = 5 somit:

$$\sigma_{\overline{x}} = \frac{0{,}064\ \text{mm}}{\sqrt{5}} \approx 0{,}02862\ \text{mm}$$

Die Berechnung der Wahrscheinlichkeit, dass die aus Stichproben vom Umfang n errechneten Mittelwerte in das Intervall [4,09 mm; 4,23 mm] fallen, erfolgt analog zu Aufgabenteil b), nur mit dem Unterschied, dass die betrachtete Zufallsgröße hier der normalverteilte Mittelwert mit den Verteilungsparametern $\mu_{\overline{x}} = 4{,}15\ \text{mm}$ und $\sigma_{\overline{x}} = 0{,}02862\ \text{mm}$ ist.

Es folgt daher:

$$z_o = \frac{x_o - \mu_{\overline{x}}}{\sigma_{\overline{x}}} \quad \Rightarrow \quad z_o = \frac{4,23 - 4,15}{0,02862} \approx 2,8$$

$$z_u = \frac{x_u - \mu_{\overline{x}}}{\sigma_{\overline{x}}} \quad \Rightarrow \quad z_u = \frac{4,09 - 4,15}{0,02862} \approx -2,1$$

Mit Hilfe der Tabelle der Summenfunktion der standardisierten Normalverteilung erhalten wir:

$$\Phi(z_o = 2,8) = 0,997445$$

$$\Phi(z_u = -2,1) = 1 - 0,982136 = 0,017864$$

Die gesuchte Differenz dieser beiden Wahrscheinlichkeiten ergibt sich somit zu:

$$\begin{aligned} P(4,09\text{ mm} \leq \overline{x}_i \leq 4,23\text{ mm}) &= \Phi(z_o) - \Phi(z_u) \\ &= 0,997445 - 0,017864 \\ &= 0,979581 \\ &\mathrel{\widehat{=}} 97,96\% \end{aligned}$$

Die Antwort auf die eingangs gestellte Frage lautet somit:

Die Wahrscheinlichkeit dafür, dass die aus Stichproben vom Umfang n = 5 errechneten Mittelwerte des Durchmessers im Bereich 4,09 mm ≤ d_i ≤ 4,23 mm liegen, beträgt rund 97,96%!