Normalverteilung

Thursday, 14 December 2023 12:33

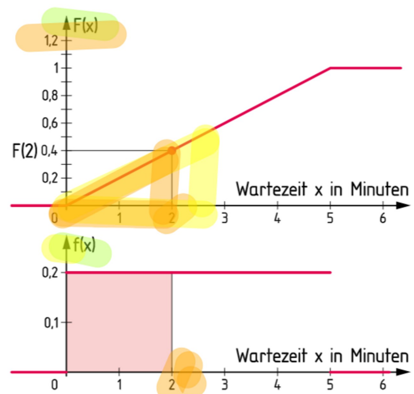
9.4 Normalverteilung

9.4.1 Grundbegriffe

Bisher wurde mit diskreten Zufallsvariablen gearbeitet, die nur bestimmte Werte annehmen. Nun werden **stetige Zufallsvariablen** behandelt, die in einem gewissen Bereich jeden beliebigen Wert annehmen können. Die zugehörige Größe ist also eine Messgröße und keine Zählgröße. Stetige Zufallsvariablen sind zum Beispiel Masse, Länge oder Wartezeit.

Da eine stetige Zufallsvariable unendlich viele Werte annehmen kann, ist es nicht möglich, für jeden Wert die Wahrscheinlichkeit anzugeben. Es gibt daher keine Wahrscheinlichkeitsfunktion wie bei diskreten Zufallsvariablen. Anhand des folgenden Beispiels wird erläutert, welche Funktion stattdessen verwendet wird:

In einer U-Bahn-Station fährt pünktlich alle 5 Minuten ein Zug ab. Die Wartezeit, mit der eine am Bahnsteig eintreffende Person rechnen muss, kann alle Werte im Intervall $[0; 5]$ annehmen und ist daher eine stetige Zufallsvariable X.



Die Wartezeit liegt mit Sicherheit zwischen 0 und 5 Minuten, also gilt: $P(X \leq 5) = 1$
Da jede Wartezeit gleich wahrscheinlich ist, steigt die Wahrscheinlichkeit für eine Wartezeit von maximal x Minuten linear an. Somit ergibt sich folgende **Verteilungsfunktion** F:
$F(x) = \frac{x}{5}$ für $0 \leq x \leq 5$

Für diskrete Zufallsvariablen kann die Verteilungsfunktion durch Aufsummieren der Einzelwahrscheinlichkeiten ermittelt werden. Diesem Vorgang entspricht im stetigen Fall das Integrieren.

Jene Funktion f, für die $F(x) = \int_{-\infty}^{x} f(\xi)\, d\xi$ gilt, heißt **Dichtefunktion** oder **Wahrscheinlichkeitsdichte**.

Fall das Integrieren.
Jene Funktion f, für die $F(x) = \int_{-\infty}^{x} f(\xi)\, d\xi$ gilt, heißt **Dichtefunktion** oder **Wahrscheinlichkeitsdichte**.

Der Graph der Dichtefunktion veranschaulicht die Verteilung einer stetigen Zufallsvariablen.
Für die Dichtefunktion f ergibt sich damit: $f(x) = \frac{1}{5}$ für $0 \leq x \leq 5$

Mithilfe der Verteilungsfunktion F kann nun die Wahrscheinlichkeit berechnet werden, **höchstens** 2 Minuten oder **zwischen** 2 Minuten und 2,5 Minuten zu warten:
$P(X \leq 2) = F(2) = 0{,}4$ $\qquad$ $P(2 \leq X \leq 2{,}5) = F(2{,}5) - F(2) = 0{,}5 - 0{,}4 = 0{,}1$

Diese Werte entsprechen jeweils dem **Flächeninhalt** unter dem Graphen der Dichtefunktion f im entsprechenden Intervall. Ein Funktionswert an der Stelle x der Dichtefunktion f entspricht aber **nicht** der Wahrscheinlichkeit $P(X = x)$. Im Bereich $[0; 5]$ liegen unendlich viele mögliche Wartezeiten, jeder einzelne Wert hat daher die Wahrscheinlichkeit „$\frac{1}{\infty}$", also 0. Die Wahrscheinlichkeit einer Wartezeit von **genau** 2 Minuten ist **null**. Das bedeutet allerdings nicht, dass die Wartezeit nie genau 2 Minuten betragen kann.

Um zu ermitteln, wie lange man im Mittel warten muss, berechnet man den **Erwartungswert**

Wahrscheinlichkeit einer Wartezeit von **genau** 2 Minuten ist **null**. Das bedeutet allerdings nicht, dass die Wartezeit nie genau 2 Minuten betragen kann.

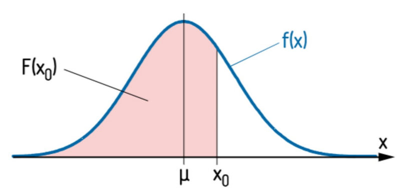
Um zu ermitteln, wie lange man im Mittel warten muss, berechnet man den **Erwartungswert**. Für stetige Zufallsvariablen wird dieser mithilfe des Integrals berechnet:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx \Rightarrow E(X) = \int_0^5 x \cdot \tfrac{1}{5}\,dx = \tfrac{x^2}{10}\Big|_0^5 = 2{,}5$$

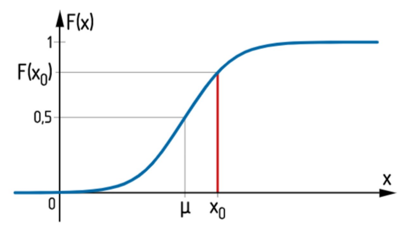
Man wartet im Mittel 2,5 Minuten.

Gaußsche Glockenkurve

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



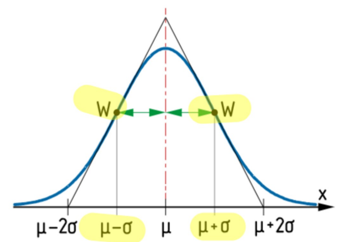
$F(x_0)$ gibt die Wahrscheinlichkeit an, mit der die normalverteilte Zufallsvariable X einen Wert von **höchstens x_0** annimmt. Sie wird durch den Flächeninhalt unter der Glockenkurve von $-\infty$ bis x_0 ermittelt. Da x_0 jeden beliebigen Wert annehmen kann, erhält man die **Verteilungsfunktion** F.



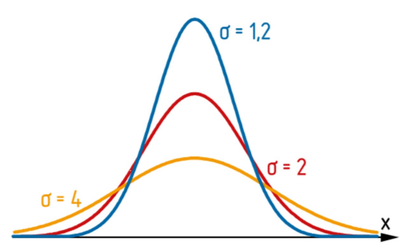
$$F(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} f(x)\,dx = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{x_0} e^{-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2}\,dx$$

Merkmale und Eigenschaften der Glockenkurve

- Der Parameter μ ist der **Erwartungswert** der Verteilung. Der Graph hat an der Stelle μ ein Maximum und ist **symmetrisch** zur Senkrechten an der Stelle μ.
- Der Parameter σ ist die **Standardabweichung** der Verteilung. Die Wendestellen des Funktionsgraphen liegen symmetrisch zum Erwartungswert bei $x = \mu \pm \sigma$. Die Wendetangenten schneiden die waagrechte Achse bei $x = \mu \pm 2\sigma$.

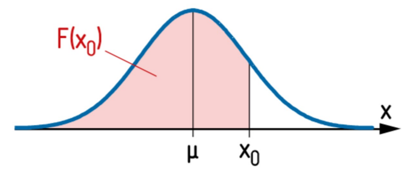


- Die Standardabweichung σ ist für die „Breite der Glocke“ bestimmend. Da die Gesamtfläche unter einer Dichtefunktion immer den Wert 1 hat, sind Glockenkurven mit kleinem σ schmäler und höher als solche mit großem σ.
- Für $x \to \pm\infty$ nähert sich der Funktionsgraph asymptotisch der waagrechten Achse.

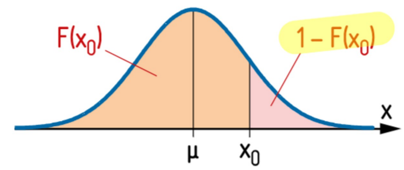


Im Folgenden werden Zusammenhänge zwischen der Fläche unter der Glockenkurve und der Wahrscheinlichkeit bei Normalverteilung gezeigt.

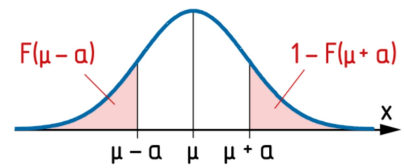
- **$P(X \leqslant x_0) = F(x_0)$**
  Die färbige Fläche („links von x_0") entspricht der Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert kleiner gleich x_0 (also höchstens x_0) annimmt.
  Da $P(X = x_0) = 0$ ist, gilt: $P(X \leqslant x_0) = P(X < x_0)$



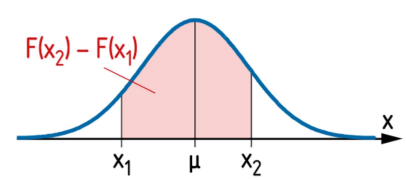
- **$P(X \geqslant x_0) = 1 - F(x_0)$**
  Die rosa unterlegte Fläche („rechts von x_0") entspricht der Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert größer gleich x_0 (also mindestens x_0) annimmt.

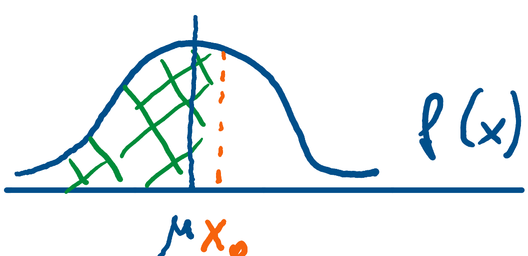


- Aufgrund der Symmetrie der Glockenkurve gilt:
  $F(\mu - a) = 1 - F(\mu + a)$

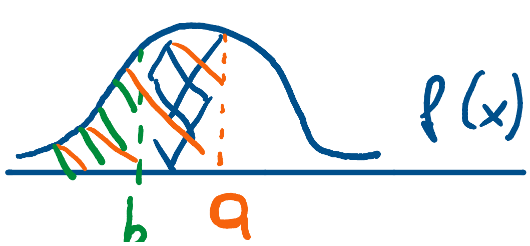


- **$P(x_1 \leqslant X \leqslant x_2) = F(x_2) - F(x_1)$**
  Die färbige Fläche entspricht der Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert zwischen x_1 und x_2 annimmt.





$$\int_{-\infty}^{x_0} f(x)\,dx$$



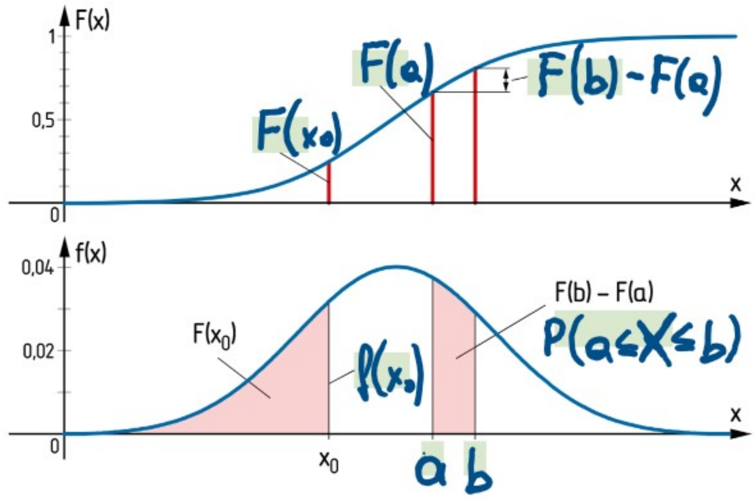
$$F(a) - F(b) = F(b \leq x \leq a)$$

9.47 In der Grafik sind die Dichtefunktion und die Verteilungsfunktion einer normalverteilten Größe dargestellt.

1) Erkläre den mathematischen Zusammenhang zwischen den beiden Funktionen.

2) Schreibe die folgenden Ausdrücke jeweils in das entsprechende Feld:

a, b, $f(x_0)$, $F(x_0)$, $F(a)$, $F(b) - F(a)$, $P(a \leq X \leq b)$



Verteilungsfunktion
$F(x)$

Dichtefunktion
$f(x)$

1) Die Verteilungsfunktion ist die Fläche zwischen dem Graphen der Dichtefunktion und der x-Achse.

$$F(x) = \int_{-\infty}^{x} f(x)\, dx$$

9.48 Ergänze die Textlücken durch Ankreuzen der richtigen Satzteile so, dass eine korrekte Aussage entsteht.

Wird der Parameter σ einer Normalverteilung ___①___, so ___②___ .

①	
verdoppelt	A
halbiert	☒
vervierfacht	C

②	
verdoppelt sich der Flächeninhalt unter der Kurve.	A
halbiert sich der Erwartungswert.	B
wird die Glockenkurve schmäler.	☒

9.4.2 Die Standardnormalverteilung – Theoretische Grundlagen

$z = \frac{x - \mu}{\sigma}$ … **Standardisierungsformel**

!