Statistik – Einführung

Stetige Zufallsvariable
Kapitel 5

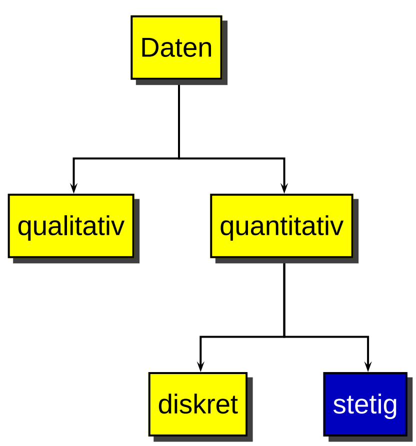
Statistik – WU Wien

Gerhard Derflinger · Michael Hauser · Jörg Lenneis · Josef Leydold ·
Günter Tirler · Rosmarie Wakolbinger

Lernziele

1. Definieren stetige Zufallsvariablen.
2. Beschreiben (stetig) gleich-, normal- und exponential-verteilte Zufallsvariablen.
3. Berechnen Erwartungswert und Varianz von stetigen ZV.
4. Bestimmen Wahrscheinlichkeiten stetiger Zufallsvariabler.

Datentypen



Stetige Zufallsvariable

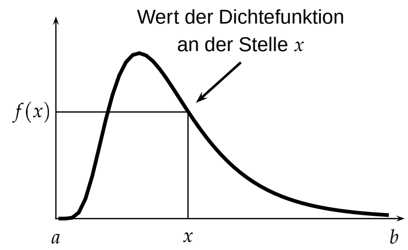
Stetige Zufallsvariable

1. Zufallsvariable
   - Das numerische Ergebnis eines Experiments.
     Gewicht eines Patienten: $\ldots, 45, 80.9, 124.4, \ldots$ – stetig
     Anzahl PKW pro Haushalt: $0, 1, 2, 3, \ldots$ – diskret
2. **Stetige** Zufallsvariable
   - Ganze Zahlen (gerundet), Bruchzahlen oder reelle Zahlen.
   - Erhält man bei Messungen.
   - Unendlich viele mögliche Werte in einem Intervall.
     Zu viele, um sie wie bei diskreten Zufallsvariablen aufschreiben zu können.

Stetige Zufallsvariable // Beispiele

Experiment	Zufallsvariable	Mögliche Werte
Wiege Personen	Gewicht (kg)	20.1, 80.5, ...
Frage nach den Ausgaben für Nahrungsmittel	Ausgaben (in €)	102, 332, ...
Messe die Zeit zwischen zwei Ankünften	Zwischenankunftszeit (min)	0, 2.4, 3.11, ...
Messe die tägliche Arbeitszeit	Stunden	6.5, 8.0, 10.2, ...

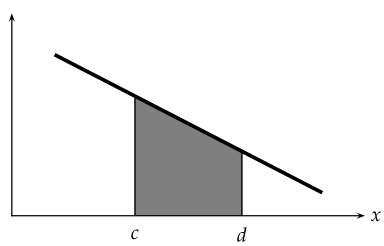
1. Mathematische Formel
2. Ein Graph, der alle
   - Werte x, und
   - Häufigkeiten $f(x)$, zeigt. $f(x)$ ist keine Wahrscheinlichkeit.
3. Eigenschaften:
   - $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1.$ Die Fläche unter der Kurve ist 1.
   - $f(x) \geq 0.$



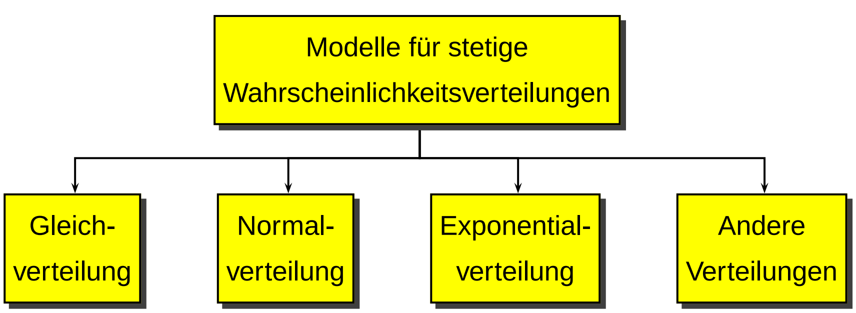
Wahrscheinlichkeit stetiger Zufallsvariabler

Die Wahrscheinlichkeit ist die Fläche unter dem Graphen der Dichtefunktion!

$$P(c \leq X \leq d) = \int_c^d f(x)\,dx$$



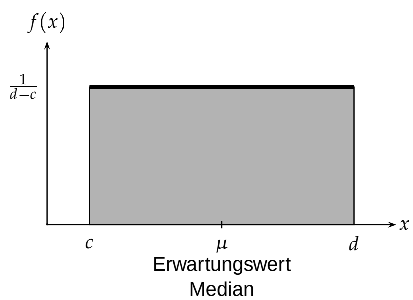
Modelle für Verteilungen stetiger Zufallsvariabler



Stetige Gleichverteilung

Stetige Gleichverteilung

1. Gleichwahrscheinliche Realisationen (Ergebnisse).
2. Dichtefunktion, $c < x < d$:
$$f(x) = \frac{1}{d-c}$$
3. Erwartungswert und Standardabweichung:
$$\mu = \frac{c+d}{2}, \qquad \sigma = \frac{d-c}{\sqrt{12}}.$$



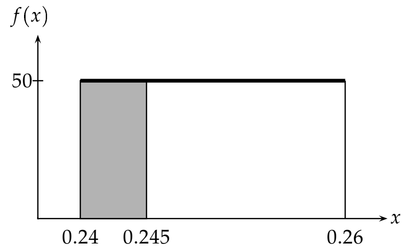
Stetige Gleichverteilung // Beispiel

Sie sind der Produktionsmanager einer Abfüllanlage für nicht alkoholische Getränke. Sie glauben, dass eine Maschine, die auf 0.25l eingestellt ist, in Wirklichkeit zwischen 0.24 und 0.26l einfüllt.

Angenommen die Abfüllmenge ist gleichverteilt über dieses Intervall. Was ist die Wahrscheinlichkeit, dass die Füllmenge eines Gebindes kleiner als 0.245l ist?

Die Zufallsvariable X heißt: Füllmenge. Sie ist stetig gleichverteilt.

$$f(x) = \frac{1}{d-c} = \frac{1}{0.26-0.24} = \frac{1}{0.02} = 50$$



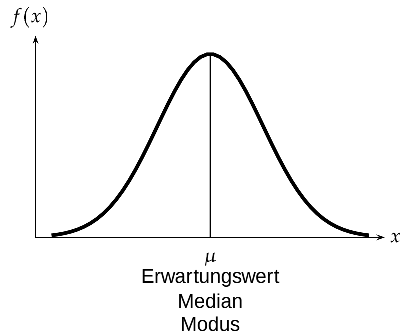
$P(0.24 \leq X \leq 0.245) =$ (Basis) $\times$ (Höhe) $=$
$= (0.245 - 0.24)(50) = 0.25$

Normalverteilung

Bedeutung der Normalverteilung

1. Einfache mathematische Eigenschaften.
2. Beschreibt gut viele stochastische (zufällige) Prozesse und stetige Phänomene.
3. Approximiert gut andere (schwierig zu berechnende) Verteilungen.
4. Ist Basis für klassische statistische Inferenz (statistisches Schließen).

1. Form: „Glockenförmig“ und symmetrisch. Gaußsche Glockenkurve
2. Erwartungswert, Median und Modus sind gleich.
3. Zufallsvariable hat eine unendliche Spannweite.



Normalverteilung // Dichtefunktion

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

$f(x)$ ... Wert der Dichte der Zufallsvariablen X an der Stelle x

μ ... Erwartungswert, Mittel der Grundgesamtheit

σ ... Standardabweichung in der Grundgesamtheit

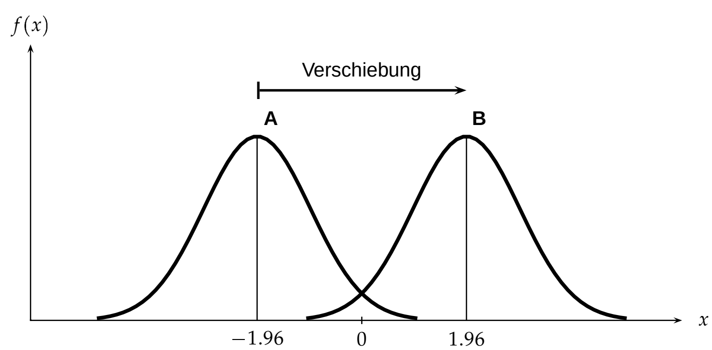
π ... $3.14159\ldots$

e ... Eulersche Zahl ($2.71828\ldots$), $\exp(x) = e^x$

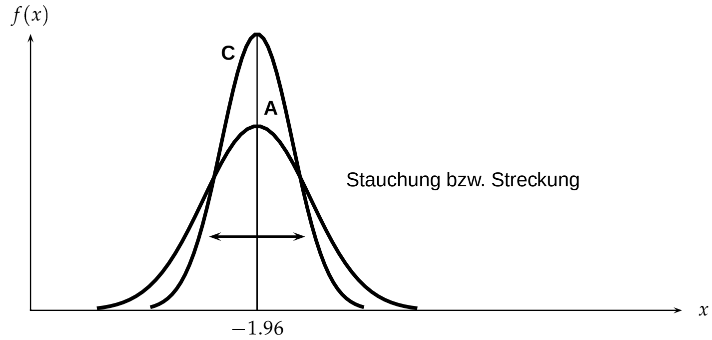
x ... Wert der Zufallsvariablen X, $(-\infty < x < \infty)$

Die Normalverteilung hat 2 Parameter: μ und σ.

Veränderung des Parameters μ



Normalverteilungen mit
A: $\mu = -1.96$, $\sigma = 1$ B: $\mu = +1.96$, $\sigma = 1$

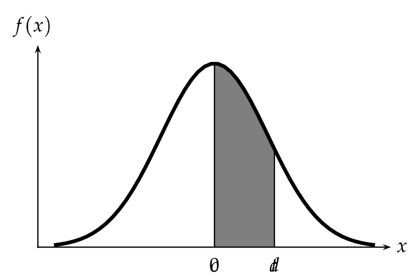


Normalverteilungen mit
A: $\mu = -1.96,\ \ \sigma = 1$
C: $\mu = -1.96,\ \ \sigma = 2/3$

Wahrscheinlichkeit und Normalverteilung

Die Wahrscheinlichkeit ist die Fläche unter dem Graphen der Dichtefunktion!

$$P(c \leq X \leq d) = \int_c^d f(x)\,dx$$



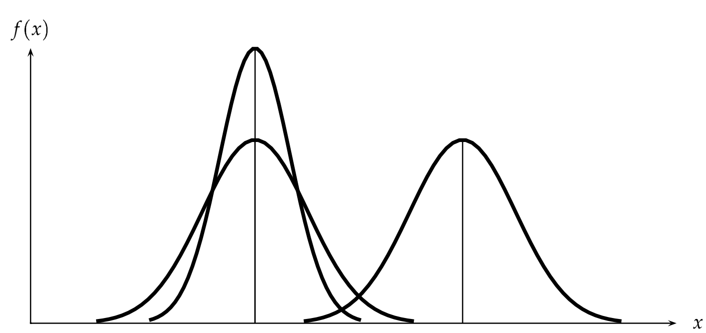
Integral nicht analytisch lösbar!
Daher sind die Werte

$$P(0 \leq X \leq t) = \int_0^t f(x)\,dx$$

tabelliert.

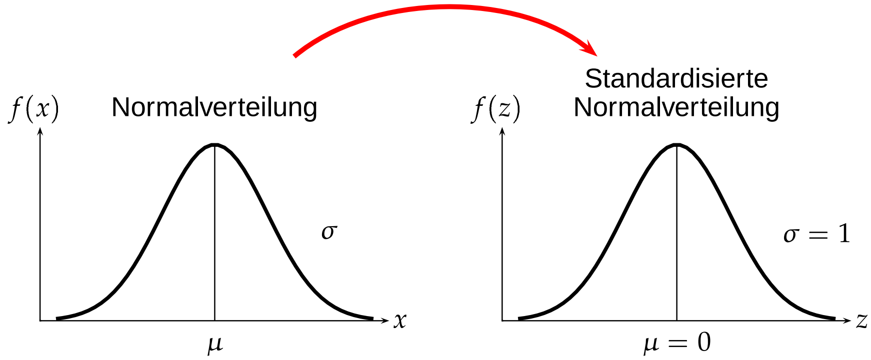
Unendliche Anzahl von Tabellen

Die Normalverteilungen unterscheiden sich in Mittel und Standardabweichung.



Jede Verteilung würde ihre eigene Tabelle benötigen.
Das sind **unendlich** viele!

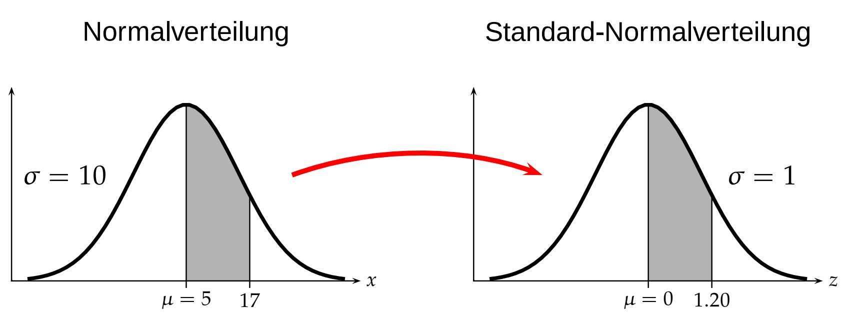
$$\boxed{Z = \frac{X-\mu}{\sigma}}$$



Nur mehr **eine** Tabelle!

Standardisieren // Beispiel

$$z = \frac{x-\mu}{\sigma} = \frac{17-5}{10} = 1.20$$



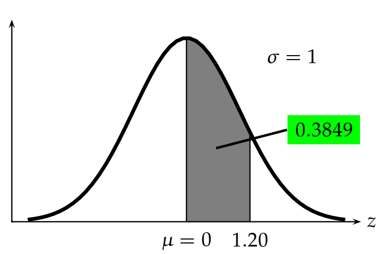
Ermitteln der Wahrscheinlichkeit

Standard-Normalverteilung
Tabelle (Auszug):

$$P(0 \leq Z \leq z)$$

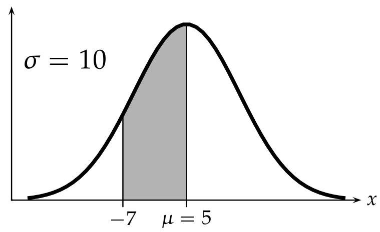
z	.00	.01	.02
0.0	.0000	.0040	.0080
0.1	.0398	.0438	.0478
0.2	.0739	.0832	.0871
...	...	...	...
1.2	.3849	.3869	.3888
1.3	.4032	.4049	.4066

Standard-Normalverteilung

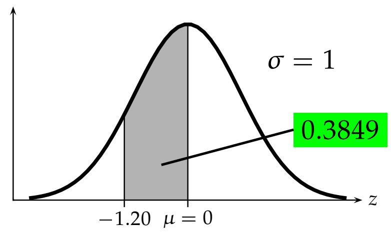


$$z = \frac{x - \mu}{\sigma} = \frac{-7 - 5}{10} = -1.20$$

Normalverteilung



Standard-Normalverteilung

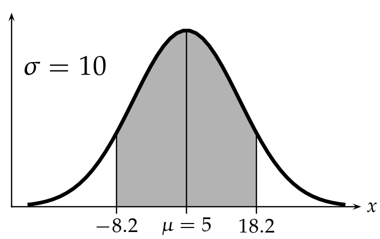


Beispiel: $P(-8.2 \le X \le 18.2) \quad (\mu = 5, \sigma = 10)$

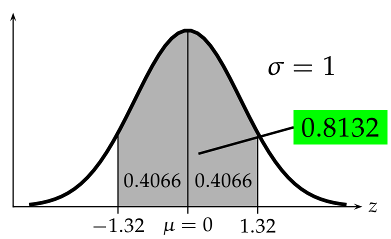
$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{-8.2 - 5}{10} = -1.32$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{18.2 - 5}{10} = 1.32$$

Normalverteilung



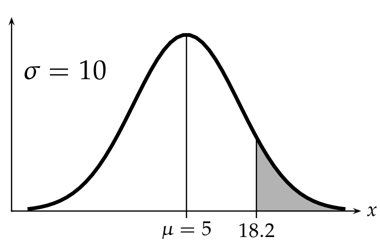
Standard-Normalverteilung



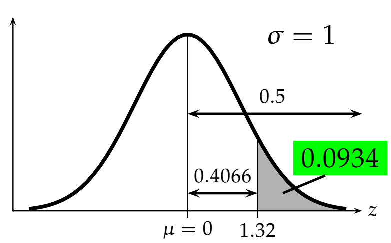
Beispiel: $P(X \ge 18.2) \quad (\mu = 5, \sigma = 10)$

$$z = \frac{x - \mu}{\sigma} = \frac{18.2 - 5}{10} = 1.32$$

Normalverteilung



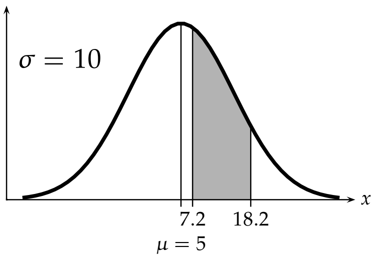
Standard-Normalverteilung



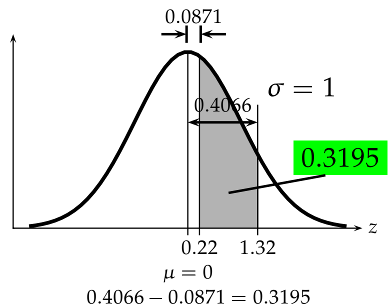
$$0.5 - 0.4066 = 0.0934$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{7.2 - 5}{10} = 0.22$$
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{18.2 - 5}{10} = 1.32$$

Normalverteilung



Standard-Normalverteilung



Normalverteilung // Beispiel

Sie arbeiten in der Qualitätskontrolle bei OSRAM. Die Lebenserwartung von Glühbirnen sei normal verteilt mit $\mu = 2000$ Stunden und $\sigma = 200$ Stunden.

Wie groß ist die Wahrscheinlichkeit, dass eine Glühbirne

A. zwischen 2000 und 2400 Stunden,

B. weniger als 1700 Stunden

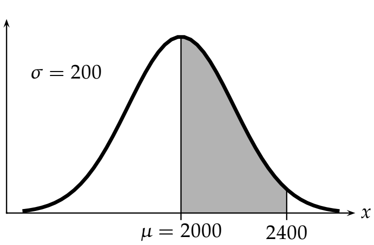
hält?

Verwenden Sie dazu die Standard-Normalverteilungstabelle aus Anhang A.

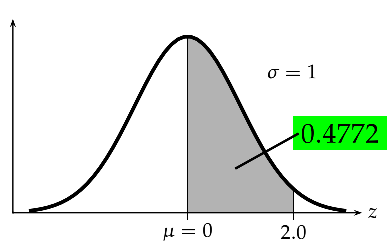
Lösung: $P(2000 \leq X \leq 2400)$

$$z = \frac{x - \mu}{\sigma} = \frac{2400 - 2000}{200} = 2.0$$

Normalverteilung

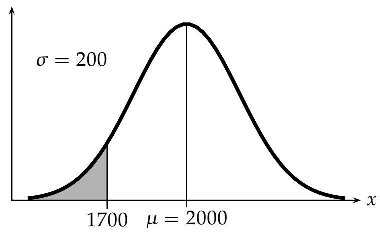


Standard-Normalverteilung

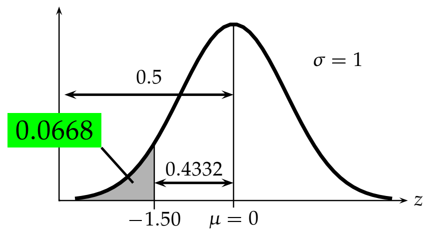


$$z = \frac{x-\mu}{\sigma} = \frac{1700-2000}{200} = -1.5$$

Normalverteilung



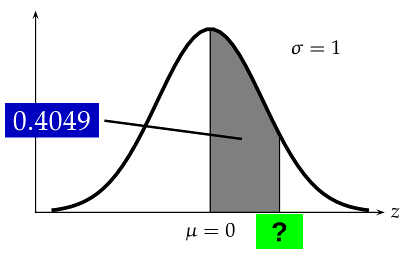
Standard-Normalverteilung



$$0.5 - 0.4332 = 0.0668$$

z-Wert bei gegebener Wahrscheinlichkeit

Wie groß ist z gegeben
$P(0 \leq Z \leq z) = 0.4049$?



$z =$ 1.31

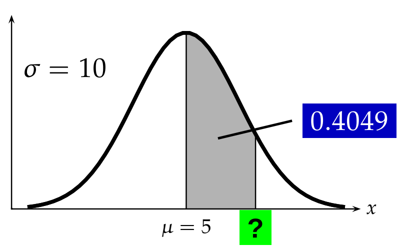
Standard-Normalverteilung
Tabelle (Auszug)

$P(0 \leq Z \leq z)$

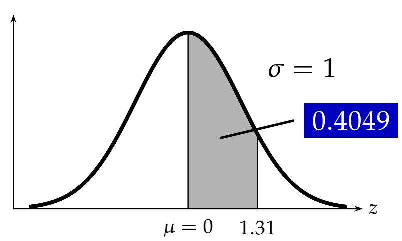
z	.00	.01	.02
0.0	.0000	.0040	.0080
0.1	.0398	.0438	.0478
0.2	.0739	.0832	.0871
...	...	...	...
1.2	.3849	.3869	.3888
1.3	.4032	0.4049	.4066

x-Wert bei gegebener Wahrscheinlichkeit

Normalverteilung



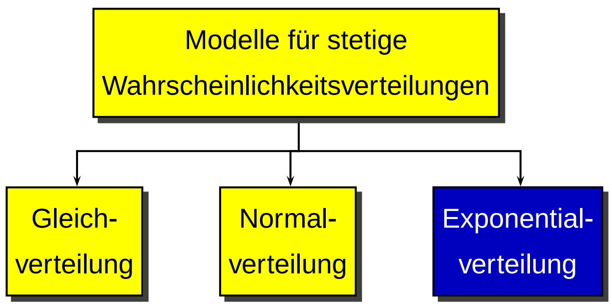
Standard-Normalverteilung



$$\boxed{x = \mu + z \cdot \sigma} = 5 + 1.31 \cdot 10 = 18.1$$

Exponentialverteilung

Modelle für Verteilungen stetiger Zufallsvariabler



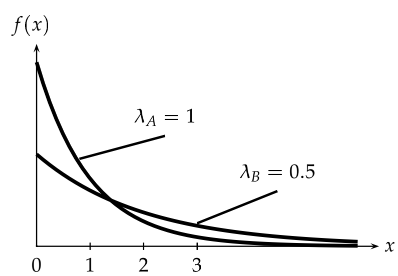
Exponentialverteilung

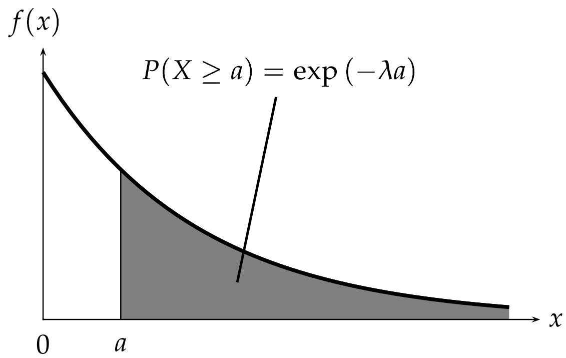
1. Beschreibt Zeit oder Distanz zwischen Ereignissen.
   - Wird bei Warteschlangen verwendet.
2. Parameter: λ
3. Dichtefunktion, $x \geq 0$:

$$\boxed{f(x) = \lambda \exp(-\lambda x)}$$

4. Erwartungswert und Standardabweichung:

$$\mu = \frac{1}{\lambda}, \qquad \sigma = \frac{1}{\lambda}.$$





Exponentialverteilung // Anwendung

Der Studiendekan hat (angenommen) eine Sekretärin. Im Durchschnitt kommt alle 5 Minuten ein Student mit einem Anliegen.

Wie groß ist die Wahrscheinlichkeit (unter der Annahme einer Exponentialverteilung), dass mehr als 15 Minuten vergehen, ohne dass ein Student ankommt?

Lösung

Die Zufallsvariable X bezeichnet die Zeit zwischen den Ankünften von zwei aufeinanderfolgenden Studenten.

$$P(X \geq a) = \exp(-\lambda a)$$

$$\mu = 5, \qquad \lambda = \frac{1}{5}, \qquad a = 15.$$

$$P(X \geq 15) = \exp\left(-\frac{1}{5} \cdot 15\right) = 0.049787 \approx 5\%$$

1. Definierten stetige Zufallsvariablen.
2. Beschrieben (stetig) gleich-, normal- und exponential-verteilte Zufallsvariablen.
3. Berechneten Erwartungswert und Varianz von stetigen ZV.
4. Bestimmten Wahrscheinlichkeiten stetiger Zufallsvariabler.