Serie 5

Aufgabe 5.1

Die Zeit, die ein Passagier an einem Flughafen Check-in Schalter verbringt ist eine Zufallsvariable mit Mittelwert 8.2 Minuten und Standardabweichung 6 Minuten. Wir beobachten zufällig 36 Passgiere.

a) Berechnen Sie die Wahrscheinlichkeit, dass die durchschnittliche Wartezeit dieser Passagiere weniger als 10 Minuten beträgt.

b) Berechnen Sie die Wahrscheinlichkeit, dass die durchschnittliche Wartezeit dieser Passagiere zwischen 5 und 10 Minuten beträgt.

c) Berechnen Sie die Wahrscheinlichkeit, dass die durchschnittliche Wartezeit dieser Passagiere mehr als 20 Minuten beträgt.

d) Alle haben wohl die Erfahrung, dass man schon länger beim Check-in. Warum ist die Wahrscheinlichkeit von d) dann so klein?

e) Gilt hier die i.i.d.-Annahme überhaupt?

Aufgabe 5.2

Das Strassenverkehrsamt hat genug Streusalz gelagert, um mit einem Schneefall von insgesamt 80 cm fertigzuwerden. Täglich fallen im Mittel 1.5 cm mit einer Standardabweichung von 0.3 cm.

Wie gross die Wahrscheinlichkeit, dass das gelagerte Salz für die nächsten 50 Tage ausreicht?

Aufgabe 5.3

Die Lebensdauer eines bestimmten elektrischen Teils ist durchschnittlich 100 Stunden mit Standardabweichung von 20 Stunden. Wir testen 16 solcher Teile.

Geben Sie die Wahrscheinlichkeit an, dass das Stichprobenmittel

a) unter 104 Stunden oder

b) zwischen 98 und 104 Stunden liegt.

Aufgabe 5.4

Ein Zigarettenhersteller gibt an, dass der Nikotingehalt in einer Zigarette durchschnittlich 2.2 mg mit Standardabweichung 0.3 mg hat. Bei einer Stichprobe von 100 zufällig ausgewählten Zigaretten ist das Stichprobenmittel allerdings 3.1 mg.

Wie hoch ist die Wahrscheinlichkeit, dass das Stichprobenmittel einen Wert von 3.1 mg oder mehr erreicht, wenn die Aussage des Zigarettenherstellers wahr ist?

Aufgabe 5.5

Ein Dozent weiss aus Erfahrung, dass bei einer Prüfung die Punktezahlen durchschnittlich bei 77 Punkte mit einer Standardabweichung von 15 Punkten liegen. In diesem Semester unterrichtet der Dozent zwei Kurse; der eine hat 25, der andere 64 TeilnehmerInnen.

a) Wie gross ist näherungsweise die Wahrscheinlichkeit, dass das durchschnittliche Prüfungsergebnis im dem Kurs mit 25 TeilnehmerInnen zwischen 72 und 82 Punkten liegt?

b) Wiederholen Sie die Rechnung aus Teil a) für den Kurs mit 64 TeilnehmerInnen.

Aufgabe 5.6

Erzeugen Sie mit **norm.rvs()** $m = 500$ Stichproben aus einer Standardnormalverteilung mit Umfang $n = 5$. Speichern Sie die Stichproben als $(n \times m)-$ Matrix ab. ***Python***-*Hinweis:*

```
import matplotlib.pyplot as plt
from scipy.stats import norm
import numpy as np

n = 5
m = 500

ran = np.array(norm.rvs(size=n*m))

sim = ran.reshape((n,m))
```

a) Stellen Sie die Stichproben als Runs graphisch dar, und staunen Sie über die zahlreichen möglichen Verläufe.

```
plt.plot(sim)
plt.show()
```

b) Berechnen Sie nun für jede Stichprobe den Mittelwert $\bar{x}$, und erzeugen Sie ein Histogramm der Mittelwerte. Aus der Theorie wissen Sie, dass die Mittelwerte normalverteilt mit Parameter $\mu = 0$ und Standardabweichung $\sigma = \frac{1}{\sqrt{n}}$ sein müssten. Überprüfen Sie dies graphisch mit Hilfe von Histogrammen und der theoretischen Wahrscheinlichkeitsdichtekurve

```
plt.hist(sim.T, bins=20, density=True, edgecolor="black",
facecolor="white")

x = np.linspace(-4, 4, num=100)
y = norm.pdf(x)

plt.plot(x,y)
plt.title("Histogramm sim")
plt.show()

sim_mean = sim.mean(axis=0)

plt.hist(sim_mean, density=True, edgecolor="black",
facecolor="white")

x = np.linspace(-4, 4, num=100)

y = norm.pdf(x, loc=0, scale=1/np.sqrt(n))

plt.plot(x,y)
plt.title("Histogramm sim_mean")

plt.show()
```

c) Wiederholen Sie die Aufgabe für $n = 2, 10, 100$.

Aufgabe 5.7

a) Eine Elektronik-Firma stellt Widerstände her, die einen Erwartungswert von $100\,\Omega$ und eine Standardabweichung von $10\,\Omega$ haben. Die Widerstände sind normalverteilt. Bestimmen Sie die Wahrscheinlichkeit, dass sich für eine zufällige Stichprobe von $n = 25$ Widerständen ein mittlerer Widerstand unter $95\,\Omega$ ergibt.

b) Nehmen Sie an, dass eine Zufallsvariable X einer uniformen Verteilung folgt, und zwar $X \sim \text{Uniform}[4, 6]$. Die Zufallsvariable $\overline{X}_{40}$ sei der Mittelwert einer i.i.d. Stichprobe vom Umfang $n = 40$. Wie lautet die Verteilung von $\overline{X}_{40}$ approximativ?

Aufgabe 5.8

Es seien $\{X_i\}_{1 \leq i \leq 50}$ unabhängige und normalverteilte Zufallsvariablen mit Erwartungswert $\mu = 1$ und Standardabweichung $\sigma = 2$. Darüber hinaus sind folgende Zufallsvariablen definiert:

$$S_n = X_1 + X_2 + \cdots + X_n$$

und

$$\overline{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \frac{S_n}{n}$$

Dabei ist $n = 50$.

a) Bestimmen Sie die Parameter der Normalverteilung von S_n sowie $\overline{X}_n$.

b) Berechnen Sie die Wahrscheinlichkeit $P\left(\mathrm{E}[X_1] - 1 \leq X_1 \leq \mathrm{E}[X_1] + 1\right)$.

c) Berechnen Sie $P\left(\mathrm{E}[S_n] - 1 \leq S_n \leq \mathrm{E}[S_n] + 1\right)$.

d) Berechnen Sie $P\left(\mathrm{E}[\overline{X}_n] - 1 \leq \overline{X}_n \leq \mathrm{E}[\overline{X}_n] + 1\right)$.

Kurzlösungen vereinzelter Aufgaben

A 5.2: ≈ 0.99

A 5.3:

a) 0.7881446

b) 0.4435663

A 5.4: ≈ 0

A 5.5:

a) 0.9044193

b) 0.9923392

A 5.7:

a) 0.0062

b) $\overline{X}_{40} \sim \mathcal{N}(5, 1/120)$

Musterlösungen zu Serie 5

Lösung 5.1

X_i ist die Zufallsvariable der Wartezeit für den i-ten Passagier. Es gilt $\mu = 8.2$ und $\sigma_X = 6$.

Wir betrachten die durchschnittliche Wartezeit $\overline{X}_{36}$, die annähernd wie folgt verteilt ist:

$$\overline{X}_{36} \sim \mathcal{N}\left(\mu, \frac{\sigma_X^2}{n}\right) = \mathcal{N}\left(8.2, \frac{6^2}{36}\right) = \mathcal{N}(8.2, 1)$$

a) Gesucht ist

$$P(\overline{X}_{36} \leq 10) = 0.9640697$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=10, loc=8.2, scale=1)

## 0.9640696808870742
```

b) Gesucht ist

$$P(5 \leq \overline{X}_{36} \leq 10) = 0.9633825$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=10, loc=8.2, scale=1) - norm.cdf(x=5, loc=8.2, scale=1)

## 0.9633825429491584
```

Der Unterschied zu b) ist sehr klein. Das heisst aber auch, dass die Wahrscheinlichkeit, dass der Durchschnitt unter 5 Minuten liegt sehr klein ist.

c) Gesucht ist

$$P(\overline{X}_{36} \geq 20) \approx 0$$

```
import numpy as np
from scipy.stats import norm

1 - norm.cdf(x=20, loc=8.2, scale=1)

## 0.0
```

Hier wird die W'keit 0 angegeben, aber dies ist sie nicht. Sie ist nur so klein, dass sie mit 0 dargestellt wird.

d) Die Wahrscheinlichkeit, dass man selber mehr als 20 Minuten warten kann, ist natürlich viel grösser.

   Die Wahrscheinlichkeit in c) beschreibt die Wahrscheinlichkeit, dass 36 zufällig beobachtete Personen durchschnittlich mehr 20 warteten und diese ist fast 0.

   Dass viele Personen durchschnittlich mehr als 20 Minuten warten müssen, ist kleiner als die Wahrscheinlichkeit dass *eine* Person mehr als 20 Minuten warten muss.

e) Diese ist hier wohl nicht gegeben.

Lösung 5.2

X_i ist die Zufallsvariable für die gefallene Menge Schnee am Tag i. Es gilt $\mu = 1.5$ und $\sigma_X = 0.3$

Wir betrachten die Schneemenge (Summe) S_{50} der nächsten 50 Tage und diese soll 80 nicht übersteigen. Es gilt annähernd

$$S_{50} \sim \mathcal{N}\left(50 \cdot \mu, 50 \cdot \sigma_X^2\right) = \mathcal{N}(75, 4.5)$$

Wir suchen

$$P(S_n \leq 80) = 0.9907889$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=80, loc=75, scale=np.sqrt(4.5))


## 0.9907889372729505
```

Lösung 5.3

X_i ist die Zufallsvariable für die Lebensdauer des Teils i. Es gilt $\mu = 100$ und $\sigma_X = 20$.

Wir betrachten die durchschnittliche Lebensdauer $\overline{X}_{16}$, die annähernd wie folgt verteilt ist:

$$\overline{X}_{16} \sim \mathcal{N}\left(\mu, \frac{\sigma_X^2}{n}\right) = \mathcal{N}\left(100, \frac{20^2}{16}\right) = \mathcal{N}(100, 25)$$

a) Gesucht ist

$$P(\overline{X}_{16} \leq 104) = 0.7881446$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=104, loc=100, scale=5)

## 0.7881446014166034
```

b) Gesucht ist

$$P(98 \leq \overline{X}_{16} \leq 104) = 0.4435663$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=104, loc=100, scale=5) - norm.cdf(x=98, loc=100, scale=5)

## 0.44356634302692755
```

Lösung 5.4

X_i ist die Zufallsvariable des Nikotingehalts in der i-ten Zigarette. Es gilt $\mu = 2.2$ und $\sigma_X = 0.3$.

Wir betrachten den durchschnittlichen Nikotingehalt $\overline{X}_{100}$, die annähernd wie folgt verteilt ist:

$$\overline{X}_{100} \sim \mathcal{N}\left(\mu, \frac{\sigma_X^2}{n}\right) = \mathcal{N}\left(2.2, \frac{0.3^2}{100}\right) = \mathcal{N}(2.2, 0.0009)$$

Gesucht ist

$$P(\overline{X}_{100} \geq 3.1) \approx 0$$

Lösung 5.5

X_i ist die Zufallsvariable für die Punktezahl der i-ten StudentIn. Es gilt $\mu = 77$ und $\sigma_X = 15$.

a) Wir betrachten die durchschnittliche Punktezahl $\overline{X}_{25}$, die annähernd wie folgt verteilt ist:

$$\overline{X}_{25} \sim \mathcal{N}\left(\mu, \frac{\sigma_X^2}{n}\right) = \mathcal{N}\left(77, \frac{15^2}{25}\right) = \mathcal{N}(77, 9)$$

Gesucht ist

$$P(72 \leq \overline{X}_{25} \leq 82) = 0.9044193$$

```
import numpy as np
from scipy.stats import norm

norm.cdf(x=82, loc=77, scale=3) - norm.cdf(x=72, loc=77, scale=3)

## 0.9044192954543706
```

b) Wir betrachten die durchschnittliche Punktezahl $\overline{X}_{64}$, die annähernd wie folgt verteilt ist:

$$\overline{X}_{64} \sim \mathcal{N}\left(\mu, \frac{\sigma_X^2}{n}\right) = \mathcal{N}\left(77, \frac{15^2}{64}\right) = \mathcal{N}(77, 225/64)$$

Gesucht ist

$$P(72 \leq \overline{X}_{64} \geq 82) = 0.9923392$$

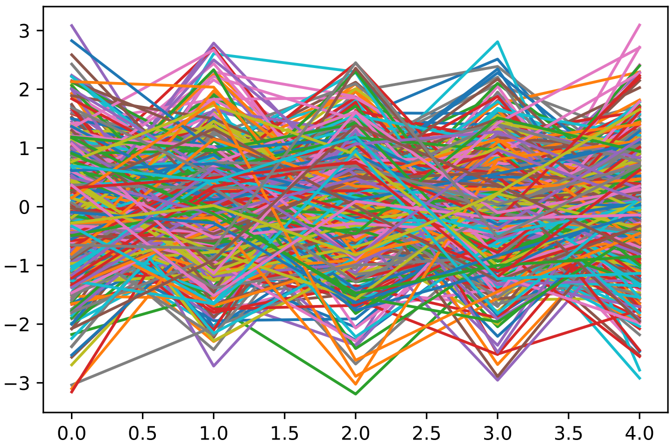
```
import numpy as np
from scipy.stats import norm

norm.cdf(x=82, loc=77, scale=np.sqrt(225/64)) - norm.cdf(x=72, loc=77, scale=np.sqrt(225/64))

## 0.9923392388648204
```

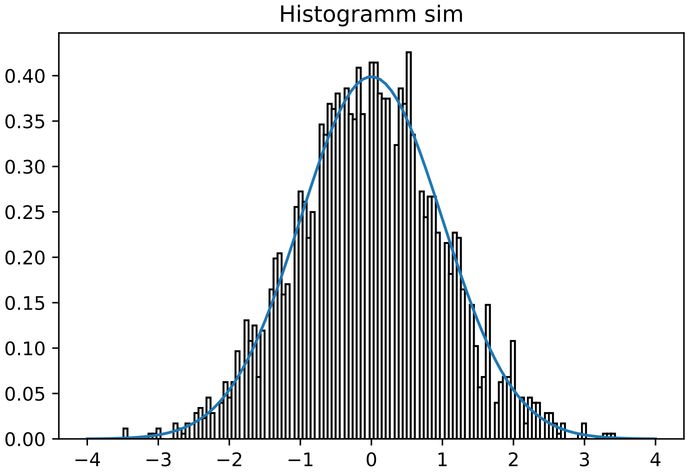
Lösung 5.6

a) (zu **R**)

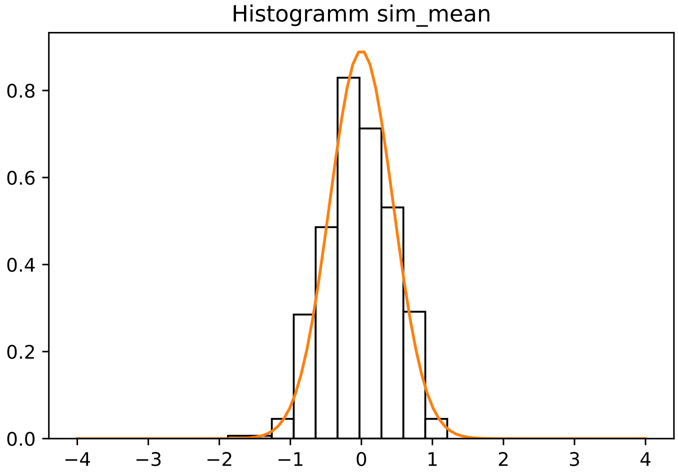


b) (zu **R**)

Für **`sim`**

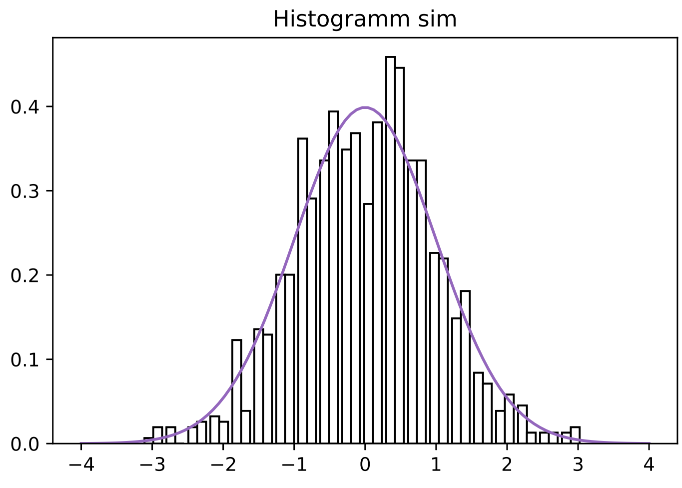


Und für **`sim_mean`**

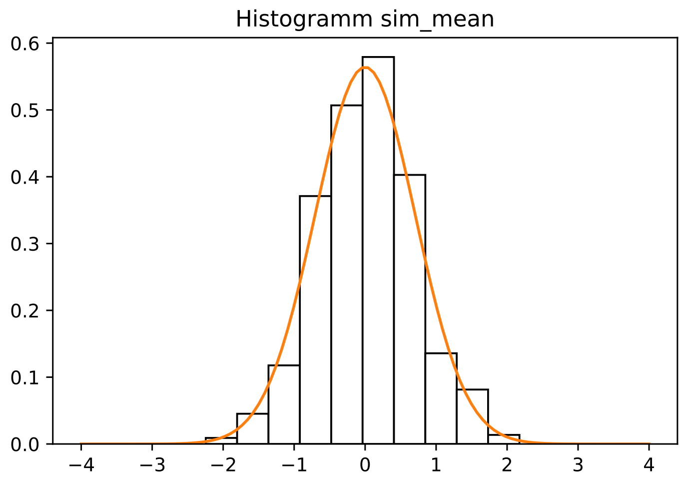


c) (zu **`R`**)

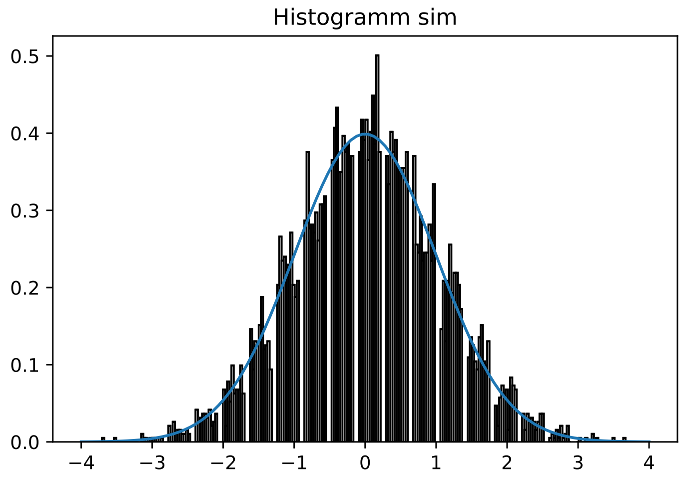
Für $n = 2$:


Histogramm sim
0.0
0.1
0.2
0.3
0.4
−4
−3
−2
−1
0
1
2
3
4

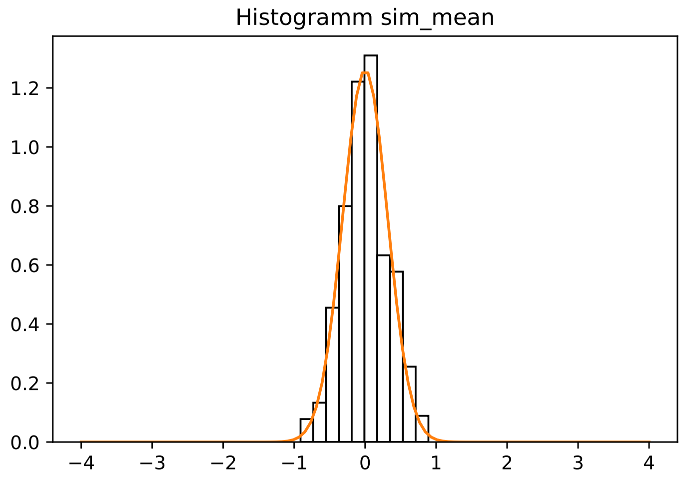
und


Histogramm sim_mean
0.0
0.1
0.2
0.3
0.4
0.5
0.6
−4
−3
−2
−1
0
1
2
3
4

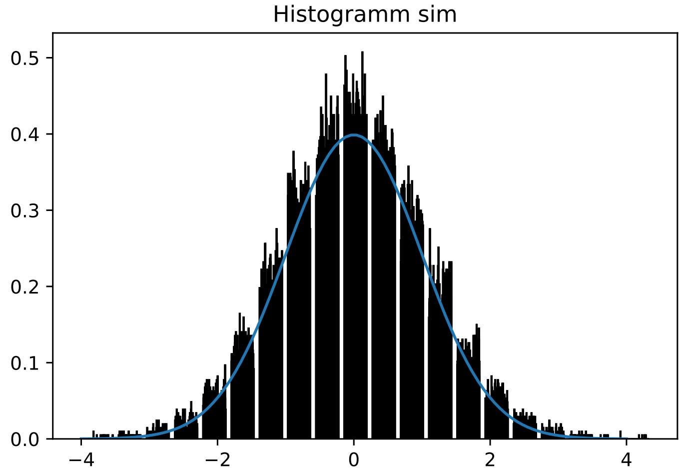
Für $n = 10$:


Histogramm sim
0.5
0.4
0.3
0.2
0.1
0.0
−4
−3
−2
−1
0
1
2
3
4

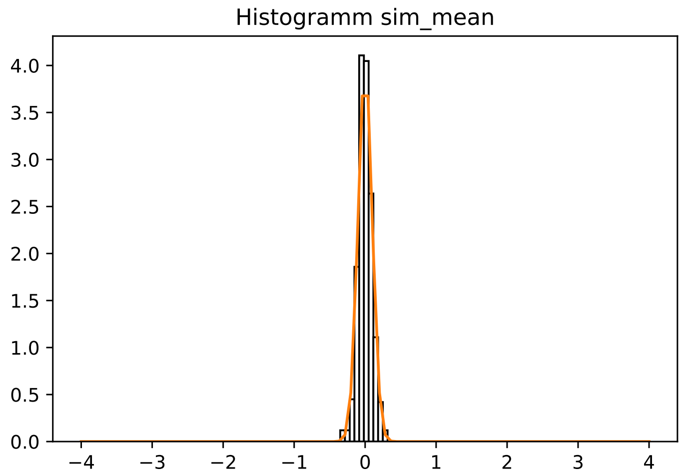
und


Histogramm sim_mean
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−4
−3
−2
−1
0
1
2
3
4

Für $n = 10$:



und



Lösung 5.7

a) Wir bezeichnen mit $\overline{X}_n$ den mittleren Widerstand in einer Stichprobe von n Widerständen, wobei $\overline{X}_n$ normalverteilt ist mit einem Mittelwert von $\mu_{\overline{X}_n} = \mu_X = 100\,\Omega$ und einer Standardabweichung von

$$\sigma_{\overline{X}_n} = \frac{\sigma_X}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2.$$

Somit ist $\overline{X}_n \sim \mathcal{N}(100, 4)$. Wir suchen nun die Wahrscheinlichkeit $P(\overline{X}_n < 95)$: (zu **R**)

```python
from scipy.stats import norm

norm.cdf(x=95, loc=100, scale=2)

## 0.006209665325776132

print(norm.cdf(x=95, loc=100, scale=2))

## 0.006209665325776132
```

b) Der Mittelwert und die Varianz von X lauten: $\mu_X = 5$ und $\sigma^2_X = (6-4)^2/12 = 1/3$. Der zentrale Grenzwertsatz deutet darauf hin, dass die Verteilung von $\overline{X}_{40}$ näherungsweise normalverteilt ist mit dem Mittelwert $\mu_{\overline{X}_{40}} = 5$ und der Varianz $\sigma_{\overline{X}_n} = \sigma_X/\sqrt{n} = 1/\sqrt{120}$. Folglich gilt

$$\overline{X}_{40} \sim \mathcal{N}(5, 1/120)$$

Lösung 5.8

a) Für S_n gilt

$$\mathrm{E}[S_n] = \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{E}[X_i] = 50 \cdot 1$$

$$\mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = 50 \cdot 4 = 200$$

S_{50} ist normal-verteilt mit $\mathcal{N}(\mu_{S_{50}} = 50, \sigma^2_{S_{50}} = 200)$. Für $\overline{X}_n$ gilt

$$\mathrm{E}[\overline{X}_n] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = 1\,, \quad \mathrm{Var}(\overline{X}_n) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = 0.08$$

$\overline{X}_{50}$ ist normal-verteilt mit $\mathcal{N}(\mu_{\overline{X}_{50}} = 1, \sigma^2_{\overline{X}_{50}} = 0.08)$.

b) X_1 ist normal-verteilt mit $\mathcal{N}(\mu_{X_1} = 1, \sigma^2_{X_1} = 4)$. Wir suchen die Wahrscheinlichkeit $P(0 \leq X_1 \leq 2)$: (zu **R**)

```python
from scipy.stats import norm

norm.cdf(x=2, loc=1, scale=2) - norm.cdf(x=0, loc=1, scale=2)

## 0.38292492254802624

print(norm.cdf(x=2, loc=1, scale=2) - norm.cdf(x=0, loc=1, scale=2))

## 0.38292492254802624
```

c) Es gilt: $S_n \sim \mathcal{N}(n\mu, n\sigma^2_X)$, also $S_{50}\mathcal{N}(50, 200)$. Wir suchen die Wahrscheinlichkeit $P(49 \leq S_{50} \leq 51)$: (zu **R**)

```
from scipy.stats import norm
import numpy as np

norm.cdf(x=51, loc=50, scale=np.sqrt(200)) - norm.cdf(x=49, loc=50, scale=np.sqrt(200))

## 0.05637197779701664

print(norm.cdf(x=51, loc=50, scale=np.sqrt(200)) - norm.cdf(x=49, loc=50, scale=np.sqrt(200)))

## 0.05637197779701664
```

d) Es gilt: $\overline{X}_n \sim \mathcal{N}(\mu, \sigma_X^2/n)$, $\overline{X}_{50} \sim \mathcal{N}(1, 4/50)$. Wir suchen die Wahrscheinlichkeit $P(0 \leq \overline{X}_{50} \leq 2)$: (zu **R**)

```
from scipy.stats import norm
import numpy as np
norm.cdf(x=2, loc=1, scale=np.sqrt(4/50))

## 0.9997965239912775

- norm.cdf(x=0, loc=1, scale=np.sqrt(4/50))

## -0.0002034760087224789

print(norm.cdf(x=2, loc=1, scale=np.sqrt(4/50)) - norm.cdf(x=0, loc=1, scale=np.sqrt(4/50)))

## 0.999593047982555
```