

Entscheidungen unter Unsicherheit

Klaus Rheinberger, FH Vorarlberg

2. Mai 2024

Inhaltsverzeichnis

1. Vorwort	3
1.1. Leitfaden	3
1.2. Literatur	3
I. Methoden	5
2. Diskrete Wahrscheinlichkeitsrechnung	6
2.1. Python	6
2.2. Wahrscheinlichkeitsräume	6
2.3. Beispiele diskreter Wahrscheinlichkeitsräume	7
2.4. Übung: Kombinatorik	7
2.5. Stochastische Unabhängigkeit	8
2.6. Übung: Binomialverteilung	8
2.7. Bedingte Wahrscheinlichkeit	9
2.8. Beispiel: (Wind-)Pocken	9
2.9. Zufallsvariablen	10
2.10. Übung: Electric Vehicle (EV) Data	12
2.11. Szenariogeneration für Simulationen	13
II. Aufgaben	18
3. Aufgaben 1	19
3.1. XXX (X Punkte)	19
4. Aufgaben 2	20
4.1. XXX (X Punkte)	20
5. Aufgaben 3	21
6. Aufgaben 4	22
7. Aufgaben 5	23
8. Aufgaben 6	24
Literaturverzeichnis	24

1. Vorwort

1.1. Leitfaden

Die Lehrveranstaltung ist in zwei Teile aufgeteilt:

- Im Teil Methoden werden die Problemstellungen, Konzepte und mathematischen an Beispielen und abstrahiert dargestellt.
- Im Teil Aufgaben werden den Studierenden Aufgaben gestellt, die sie selbstständig lösen und präsentieren sollen.

1.2. Literatur

- Eisenführ, Franz (2010): Rationales Entscheiden. 5. Auflage, 2010, Springer.
- Nyberg, Svein Olav (2018): The Bayesian Way: Introductory Statistics for Economists and Engineers. 2018, Wiley-Blackwell.
- Bronson, Richard; Naadimuthu, Govindasami (1997): Schaum's Outline of Operations Research. 2. Aufl., McGraw-Hill Professional.
- Schiller, John J.; Srinivasan, R. Alu; Spiegel, Murray R. (2013): Schaum's Outline of Probability and Statistics. 4 ed., McGraw-Hill Education.
- Bertsimas, Dimitris (2004): Data, Models, and Decisions: The Fundamentals of Management Science. Belmont, Mass: Dynamic Ideas Llc.: **Kapitel 1**
- Peterson, Martin (2017): An Introduction to Decision Theory. 2. Edition, Cambridge University Press.
- Cornuéjols, Gérard; Peña, Javier; Tütüncü, Reha (2018): Optimization Methods in Finance. 2. Aufl., Cambridge University Press.
- Martin, Osvaldo (2018): Bayesian Analysis with Python: Introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ, 2nd Revised edition, Packt Publishing.
- Kovacevic, Raimund M.; Pflug, Georg Ch; Vespucci, Maria Teresa (2013): Handbook of Risk Management in Energy Production and Trading. Springer.
- Birge, John R.; Louveaux, François (2011): Introduction to Stochastic Programming. 2nd ed. 2011. New York: Springer.
- Jordaan, Ian (2011): Decisions under Uncertainty: Probabilistic Analysis for Engineering Decisions. 1st edition, Cambridge University Press.
- Bishop, Christopher (2007): Pattern Recognition and Machine Learning (Information Science and Statistics). 2nd printing 2011, Springer.
- Kall, Peter; Mayer, János (2012): Stochastic Linear Programming: Models, Theory, and Computation. Softcover reprint of hardcover 2nd ed. 2011. Springer.
- King, Alan J. J.; Wallace, Stein W. (2014): Modeling with Stochastic Programming. 2012th Ed., Springer.
- Shapiro, Alexander; Philpott, Andy (2007): [A Tutorial on Stochastic Programming](#).
- Stone, James V. (2016): Bayes' Rule With Python: A Tutorial Introduction to Bayesian Analysis. First Edition. Sebtel Press.
- Wallace, Stein W.; Ziemba, William T. (2005): Applications of Stochastic Programming. Society for Industrial and Applied Mathematics.
- Bosch, Karl (2011): Elementare Einführung in die Wahrscheinlichkeitsrechnung: Mit 82 Beispielen und 73 Übungsaufgaben mit vollständigem Lösungsweg. 11., akt. Aufl., Vieweg+Teubner Verlag.
- Cramer, Erhard; Kamps, Udo (2020): Grundlagen der Wahrscheinlichkeitsrechnung und Statistik: Eine Einführung für Studierende der Informatik, der Ingenieur- und Wirtschaftswissenschaften. 5., erw. u. korr. Aufl., Springer Spektrum.

- Luenberger, David G. (2009): Investment Science: International Edition, Oxford University Press.
- Wasserman, Larry (2004): All of Statistics: A Concise Course in Statistical Inference. Corr. 2nd printing 2004. Springer.
- Diwekar, Urmila (2020): Introduction to Applied Optimization. 3. Aufl., Springer.
- Mondello, Enzo (2015): Portfoliomanagement: Theorie und Anwendungsbeispiele. 2. Auflage, Springer. [PDF-Download](#)
- Jason, Brownlee (2021): Probability for Machine Learning. [URL](#)
- Williams, H. Paul (2013): Model Building in Mathematical Programming. 5th edition, Wiley.

Teil I.

Methoden

2. Diskrete Wahrscheinlichkeitsrechnung

2.1. Python

```
import numpy as np
import math
import matplotlib.pyplot as plt
import pandas as pd
```

2.2. Wahrscheinlichkeitsräume

Wahrscheinlichkeitsräume dienen der Beschreibung von Unsicherheit, Unwissenheit, Zufall, etc.. [Kolmogorow](#) hat in den 1930er Jahren die Wahrscheinlichkeitstheorie axiomatisch begründet, siehe z. B. [Wikipedia: Wahrscheinlichkeitstheorie](#).

Die Menge aller Möglichkeiten wird durch den **Ereignisraum** Ω beschrieben. Ein Element ω des Ereignisraums heißt **Ergebnis**. Eine Teilmenge A des Ereignisraums heißt **Ereignis**. Ein Ereignis, das genau ein Element besitzt, heißt **Elementarereignis**.

Beispiel einfacher Würfelwurf: $\Omega = \{\text{Zahl 1, Zahl 2, Zahl 3, Zahl 4, Zahl 5, Zahl 6}\}$. Die Elementarereignisse sind $\{\text{Zahl 1}\}, \dots, \{\text{Zahl 6}\}$. Das Ereignis “Es fällt eine gerade Zahl” entspricht der Teilmenge $A = \{\text{Zahl 2, Zahl 4, Zahl 6}\}$. Die Bezeichnung der Ergebnisse ist nicht ausschlaggebend und hätte man auch anders wählen können.

Einige abgeleitete Begriffe:

- Das Ereignis $A \cap B$ (sprich **A und B**, genannt der **Durchschnitt** von A und B) tritt genau dann ein, wenn sowohl A als auch B , wenn also beide eintreten.
- Das Ereignis $A \cup B$ (sprich **A oder B**, genannt **Vereinigung** von A und B) tritt genau dann ein, wenn A oder B oder beide eintreten, wenn also mindestens eines der Ereignisse A, B eintritt.
- Das Ereignis \bar{A} (sprich **A nicht**, oder **nicht A**) tritt genau dann ein, wenn das Ereignis A nicht eintritt. Man nennt \bar{A} das zu A entgegengesetzte Ereignis oder das **Komplementärereignis** von A .
- Das Ereignis $A \setminus B$ (sprich **A ohne B**) tritt genau dann ein, wenn A eintritt und B nicht. $A \setminus B$ heißt die **Differenz** von A und B .
- Zwei Ereignisse A und B heißen **inkompatibel** (oder unvereinbar oder unverträglich oder disjunkt), wenn sie nicht beide gleichzeitig eintreten können, wenn also gilt $A \cap B = \emptyset$.

Bemerkung: Die leere Menge \emptyset beschreibt das “Null-Ergebnis”, das nie eintritt.

Axiome von Kolmogorow: Hier für diskrete (=abzählbare) Ereignisräume. Für kontinuierliche Ereignisräume muss die Menge der Ereignisse präzisiert werden.

1. Für jedes Ereignis $A \subseteq \Omega$ ist die Wahrscheinlichkeit von A eine reelle Zahl zwischen 0 und 1: $0 \leq P(A) \leq 1$
2. Das sichere Ereignis Ω hat die Wahrscheinlichkeit 1: $P(\Omega) = 1$.
3. Die Wahrscheinlichkeit der Vereinigung inkompatibler Ereignisse ist gleich der Summe der Wahrscheinlichkeiten der einzelnen Ereignisse: Für $A \cap B = \emptyset$ gilt $P(A \cup B) = P(A) + P(B)$.

Folgerung: $P(\emptyset) = 0$.

Der Ereignisraum Ω zusammen mit der Wahrscheinlichkeitsabbildung P heißt Wahrscheinlichkeitsraum.

2.3. Beispiele diskreter Wahrscheinlichkeitsräume

Hochwasser: Siehe Übungsbeispiel Versicherungsschutz im Abschnitt Entscheidungsanalyse, $\Omega = \{HH, HK, KH, KK\}$ mit

$$P(HH) = 0.01 \quad (2.1)$$

$$P(HK) = 0.09 \quad (2.2)$$

$$P(KH) = 0.09 \quad (2.3)$$

$$P(KK) = 0.81 \quad (2.4)$$

Laplace-Experimente: Wenn man im Falle endlich vieler Elementarereignisse annimmt, dass alle mit der gleichen Wahrscheinlichkeit eintreten, so spricht man von einem Laplace-Experiment. Die Mächtigkeit einer endlichen Menge M ist die Anzahl ihrer Elemente und wird als $|M|$ geschrieben. Für einen Ereignisraum Ω mit $|\Omega| = n$, der ein Laplace-Experiment beschreibt, gilt somit

$$P(\omega) = \frac{1}{n} \quad \text{für alle } \omega \in \Omega.$$

Beispiele:

- ideale Münze: $P(\omega) = \frac{1}{2}$
- idealer Würfel: $P(\omega) = \frac{1}{6}$

Folgerung: Ist A ein Ereignis der Mächtigkeit $|A| = m$, so ist A die Vereinigung von m Elementarereignissen. Jedes davon hat die Wahrscheinlichkeit $\frac{1}{n}$, also ist $P(A) = m \frac{1}{n} = \frac{m}{n}$. Man erhält so den Zusammenhang

$$P(A) = \frac{|A|}{|\Omega|}.$$

Bei Laplace-Experimenten ist die Wahrscheinlichkeit eines Ereignisses also gleich der Zahl der für dieses Ereignis günstigen Ergebnisse, dividiert durch die Zahl der insgesamt möglichen Ergebnisse.

2.4. Übung: Kombinatorik

Erinnerung: n verschiedene Dinge lassen sich unter Berücksichtigung der Reihenfolge auf

$$n! = 1 \cdot 2 \cdot \dots \cdot (n-1) \cdot n$$

(sprich “ n Fakultät”, in Österreich auch **Faktorielle** genannt) verschiedene Arten anordnen.

```
# compute the factorial in python with the math package:  
n = 4  
math.factorial(n) # compare with 4*3*2*1
```

24

Problemstellung: Bei einer Geburtstagsfeier sollen zehn Personen an einem runden Tisch Platz nehmen. Die Tischordnung wird zufällig ausgelost. Herr Meyer möchte gerne neben Fr. Schultze sitzen. Man berechne die Wahrscheinlichkeit, mit der dieses Ereignis eintritt.

2.5. Stochastische Unabhängigkeit

Die Ereignisse A und B heißen **stochastisch unabhängig** oder kürzer **unabhängig**, falls gilt:

$$P(A \cap B) = P(A) \cdot P(B).$$

Beispiel: Werfen von zwei Würfeln

Wir modellieren die Ergebnisse des Würfels von zwei Würfeln als Laplace-Experiment: Es gibt 36 Elementarereignisse vom Typ (Z_1, Z_2) , wobei Z_1 das Ergebnis des ersten Würfels und Z_2 das Ergebnis des zweiten Würfels angibt. Wir erhalten: $P(Z_1, Z_2) = \frac{1}{36}$.

- Mit $(Z_1, *)$ bezeichnen wir das Ereignis, dass der erste Würfel die Zahl Z_1 zeigt, und
- mit $(*, Z_2)$ bezeichnen wir das Ereignis, dass der zweite Würfel die Zahl Z_2 zeigt.

Es gilt: $P(Z_1, *) = \frac{6}{36} = \frac{1}{6} = P(*, Z_2)$ und somit $P(Z_1, Z_2) = P(Z_1, *) \cdot P(*, Z_2)$.

2.6. Übung: Binomialverteilung

Der **Binomialkoeffizient** $\binom{n}{k}$ gibt an, auf wie viele verschiedene Arten man k bestimmte Objekte aus einer Menge von n verschiedenen Objekten auswählen kann - ohne Zurücklegen und ohne Beachtung der Reihenfolge. Der Binomialkoeffizient ist also die Anzahl der k -elementigen Teilmengen einer n -elementigen Menge. Berechnung:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

```
# compute binomial coefficients in python with the math package:  
n = 10  
k = 3  
math.comb(n, k)
```

120

Binomialverteilung: Sie beschreibt die Anzahl der Erfolge in einer Serie von gleichartigen und unabhängigen Versuchen, die jeweils genau die zwei möglichen Ergebnisse Erfolg und Nicht-Erfolg haben. Solche Versuchsserien werden auch Bernoulli-Prozesse genannt. Ist p die Erfolgswahrscheinlichkeit bei einem Versuch und n die Anzahl der Versuche, dann bezeichnet man mit $B(k|p, n)$ die Wahrscheinlichkeit, genau k Erfolge zu erzielen. Berechnung:

$$B(k|p, n) = \binom{n}{k} p^k (1-p)^{n-k}$$

für $k \in \{0, \dots, n\}$ und Null sonst.

Aufgaben:

1. Begründen Sie die Formel für die Binomialkoeffizienten.
2. Begründen Sie die Formel für die Binomialverteilung.
3. Stellen Sie die Binomialverteilungen für $p = 0.5$ und $n = 5, 10$ und 100 grafisch dar. Verwenden Sie die Funktion `binom.pmf()` aus dem Modul `scipy.stats`.
4. Die Wahrscheinlichkeit, daß eine an einer bestimmten Krankheit leidende Person durch ein bestimmtes Medikament geheilt werde, sei 0.8 . Das Medikament werde 10 Patienten verabreicht. Mit welcher Wahrscheinlichkeit werden mindestens 8 der 10 Patienten geheilt? Dabei sei vorausgesetzt, daß die Heilerfolge für die einzelnen Patienten voneinander unabhängig sind und die Heilwahrscheinlichkeit bei allen Personen gleich 0.8 ist.

2.7. Bedingte Wahrscheinlichkeit

Das Konzept der bedingten Wahrscheinlichkeit, siehe auch [Wikipedia: Bedingten Wahrscheinlichkeit](#), dient der Beschreibung des Einflusses von Vor- oder Zusatzinformationen. Unter einer bedingten Wahrscheinlichkeit versteht man die Wahrscheinlichkeit für das Eintreten eines Ereignisses A unter der Voraussetzung, dass das Eintreten eines anderen Ereignisses B bereits bekannt ist. Natürlich muss B eintreten können, es darf also kein unmögliches Ereignis sein, d. h. $P(B) > 0$. Dann heißt die durch

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

definierte Zahl $P(A|B)$ die **bedingte Wahrscheinlichkeit** von A unter der Bedingung B , nämlich dass das Eintreten des Ereignisses B bereits bekannt ist, kurz: **die bedingte Wahrscheinlichkeit von A unter B** . Die bedingte Wahrscheinlichkeit von A unter B misst den Wahrscheinlichkeitsanteil von $A \cap B$ an B .

Wenn die Ereignisse A und B **stochastisch unabhängig** sind, gilt:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A).$$

Beispiel Würfelwurf: Wie groß ist die Wahrscheinlichkeit für das Auftreten der Zwei unter der Bedingung, dass eine gerade Zahl auftritt? Antwort: $1/3$

2.8. Beispiel: (Wind-)Pocken

- [Wikipedia: Windpocken](#): Kinderkrankheit, englisch: chickenpox.
- [Wikipedia: Pocken](#): lebensbedrohliche Infektionskrankheit, englisch: smallpox. “Am 8. Mai 1980 wurde von der WHO festgestellt, dass die Pocken ausgerottet sind. Für das weltweite Eliminationsprogramm wurden 2,4 Milliarden Impfdosen verabreicht [...]. Dennoch sind weitere Pockeninfektionen nicht völlig ausgeschlossen. Offiziell existiert das Virus noch in zwei Laboratorien der Welt; [...]”. Siehe insbesondere [Pockengefahr nach der Ausrottung](#).

Beide Krankheiten haben u. a. Flecken als Symptome, siehe die Bilder auf Wikipedia. Wir untersuchen folgende hypothetische Situation: Sie haben an Ihrem Körper Flecken und gehen zu Ihrer Hausärztin, die Ihnen folgendes sagt: “90 % der Menschen, die die Pocken haben, haben die gleichen Symptome wie Sie.” Sie bekommen Angst. Sie fragen die Ärztin: “Ja, aber wie hoch ist die Wahrscheinlichkeit, dass ich die Pocken habe, wenn ich diese Symptome habe?”. “Ah”, sagt sie, “eine sehr gute Frage.” Nachdem sie einige Rechnungen auf einen Zettel gekritzelt hat, antwortet sie auf Ihre Frage mit: “Die Wahrscheinlichkeit, dass Sie die Pocken haben, wenn Sie diese Symptome haben, beträgt (nur) 1,1 %.” Sie sind (etwas) erleichtert und fragen sich: “Was hat die Ärztin da gerade gerechnet?”

Quelle: James V. Stone: Bayes’ Rule With Python. Sebtel Press, 2016.

Der Satz von Bayes:

Die gemeinsame Wahrscheinlichkeit $P(A \cap B) = P(B \cap A)$ der Ereignisse A und B , kurz die Wahrscheinlichkeit von A und B , lässt sich mit Hilfe der bedingten Wahrscheinlichkeiten $P(A|B)$ und $P(B|A)$ auf zwei Arten berechnen:

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \\ P(B \cap A) &= P(B|A)P(A) \end{aligned}$$

Es folgt der **Satz von Bayes**:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Siehe auch [Wikipedia: Satz von Bayes](#).

Berechnungen zum Beispiel:

Die Datenlage der Ärztin umfasst folgende Ereignisse und Wahrscheinlichkeiten:

- $P(\text{Flecken}|\text{Pocken}) = 0.9$
- $P(\text{Flecken}|\text{Windpocken}) = 0.8$
- $P(\text{Pocken}) = 0.001$
- $P(\text{Windpocken}) = 0.1$
- $P(\text{Flecken}) = 0.081$

Anwendung des Satzes von Bayes:

$$\begin{aligned}P(\text{Pocken}|\text{Flecken}) &= \frac{P(\text{Flecken}|\text{Pocken})P(\text{Pocken})}{P(\text{Flecken})} \\ &= \frac{0.9 \times 0.001}{0.081} \\ &= 0.011\end{aligned}$$

```
print(f"P(Pocken|Flecken) = {(0.9*0.001)/0.081:.3f}")
```

$P(\text{Pocken}|\text{Flecken}) = 0.011$

$$\begin{aligned}P(\text{Windpocken}|\text{Flecken}) &= \frac{P(\text{Flecken}|\text{Windpocken})P(\text{Windpocken})}{P(\text{Flecken})} \\ &= \frac{0.8 \times 0.1}{0.081} \\ &= 0.988\end{aligned}$$

```
print(f"P(Windpocken|Flecken) = {(0.8*0.1)/0.081:.3f}")
```

$P(\text{Windpocken}|\text{Flecken}) = 0.988$

2.9. Zufallsvariablen

... dienen der Beschreibung von unsicheren (Mess)Größen.

Definition: Eine **Zufallsvariable** X ist eine Funktion auf einem Wahrscheinlichkeitsraum Ω , d. h. $X : \Omega \rightarrow \mathbb{R}$.

Schreibweisen: Die Wahrscheinlichkeit, dass eine Zufallsvariable X einen bestimmten Wert $x_0 \in \mathbb{R}$ annimmt, wird mit $P(X = x_0)$ geschrieben. Analog schreibt man $P(X \in A)$ für die Wahrscheinlichkeit, dass X einen Wert in der Menge A annimmt, und $P(X \leq x_0)$ für die Wahrscheinlichkeit, dass X einen Wert kleiner gleich x_0 annimmt.

Beispiele:

- **Hochwasser:** Siehe Übungsbeispiel Versicherungsschutz im Abschnitt Entscheidungsanalyse. Zum Beispiel die Strategie Übersiedeln entspricht der Zufallsvariablen $X(HH) = 4$, $X(HK) = 3$, $X(KH) = 3$, $X(KK) = 2$.

Vers., Vers.	HW HW	HW kHW	kHW HW	kHW kHW	EW	Std
Wahrscheinlichkeit	1,00%	9,00%	9,00%	81,00%		
Zielgröße	2,3	2,3	2,3	2,3	2,3	0,00000

Vers., nicht Vers.	HW HW	HW kHW	kHW HW	kHW kHW	EW	Std
Wahrscheinlichkeit	1,00%	9,00%	9,00%	81,00%		
Zielgröße	21,1	1,1	21,1	1,1	3,1	6,00000

nicht. Vers., Vers.	HW HW	HW kHW	kHW HW	kHW kHW	EW	Std
Wahrscheinlichkeit	1,00%	9,00%	9,00%	81,00%		
Zielgröße	21,2	21,2	1,2	1,2	3,2	6,00000

nicht Vers., nicht Vers.	HW HW	HW kHW	kHW HW	kHW kHW	EW	Std
Wahrscheinlichkeit	1,00%	9,00%	9,00%	81,00%		
Zielgröße	40	20	20	0	4	8,48528

Übersiedeln	HW HW	HW kHW	kHW HW	kHW kHW	EW	Std
Wahrscheinlichkeit	1,00%	9,00%	9,00%	81,00%		
Zielgröße	4	3	3	2	2,2	0,42426

Abbildung 2.1.: EA_Uebung_Versicherungsschutz_Matrix.png

- **Würfelspiel:** Bei Zahlen 1 und 2 bezahlt man jeweils 1 EUR, bei der Zahl 3 erhält man 3 EUR, und bei den restliche Zahlen passiert nichts: $X(\text{Zahl } 1) = X(\text{Zahl } 2) = -1$, $X(\text{Zahl } 3) = 3$, $X(\text{Zahl } 4) = X(\text{Zahl } 5) = X(\text{Zahl } 6) = 0$. Mit X wird der Geldbetrag bezeichnet. Dann gilt $P(X = 0) = \frac{3}{6}$, $P(X = -1) = \frac{2}{6}$, $P(X = 3) = \frac{1}{6}$

Begriffe und Kennzahlen:

- Die Gesamtheit der Zahlenpaare (x_i, p_i) mit x_i den möglichen Werten von X und $p_i = P(X = x_i)$ heißt **Verteilung** der diskreten Zufallsvariablen X .
- Der **Erwartungswert** μ der Zufallsvariablen X ist definiert als

$$\mu = \mathbb{E}(X) = \sum_i x_i P(X = x_i) = \sum_i x_i p_i.$$

- Die **Varianz** σ^2 der Zufallsvariablen X ist definiert als

$$\sigma^2 = \text{Var}(X) = \mathbb{E}([X - \mu]^2) = \sum_i (x_i - \mu)^2 P(X = x_i) = \sum_i (x_i - \mu)^2 p_i.$$

- Die **Standardabweichung** $\text{Std}(X) = \sigma$ der Zufallsvariablen X ist definiert als die Wurzel der Varianz.

Darstellungen:

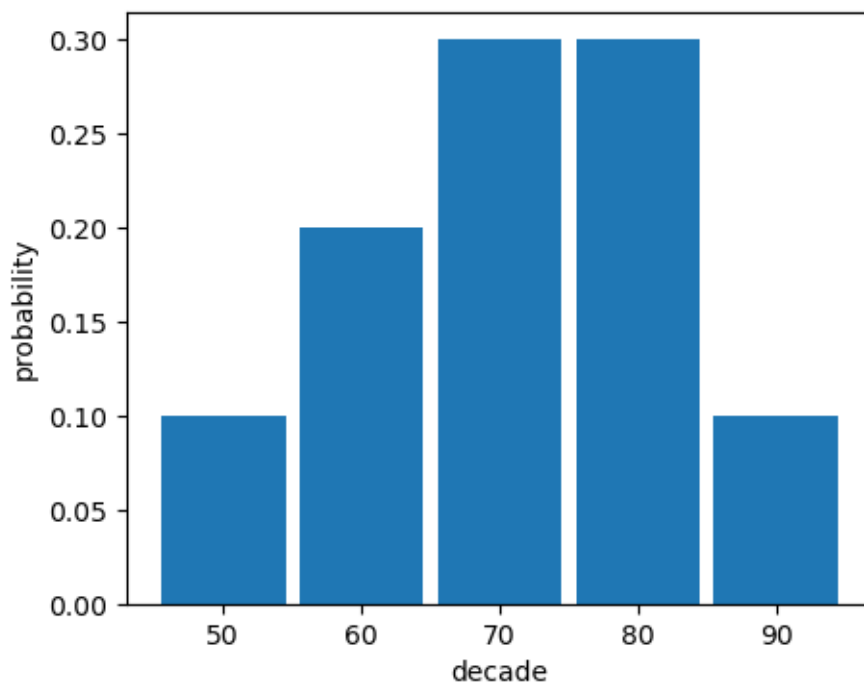
Eine diskrete Zufallsvariable wird typischerweise nicht als Funktion $X : \Omega \rightarrow \mathbb{R}$ angegeben, sondern "nur" über die Zahlenpaare (x_i, p_i) mit x_i den möglichen Werten von X und $p_i = P(X = x_i)$. Diese können z. B. tabellarisch oder als Histogramm dargestellt werden:

```
# Example: life expectancy in decades
x = np.array([50, 60, 70, 80, 90])
p = np.array([0.1, 0.2, 0.3, 0.3, 0.1])
```

```
print(f"decade | probability")
print("-----")
for x_k, p_k in zip(x, p):
    print(f"  {x_k} | {p_k}")
```

```
decade | probability
-----
    50 | 0.1
    60 | 0.2
    70 | 0.3
    80 | 0.3
    90 | 0.1
```

```
plt.figure(figsize=(5, 4))
plt.bar(x, p, width=9)
plt.xticks(x)
plt.xlabel('decade')
plt.ylabel('probability');
```



2.10. Übung: Electric Vehicle (EV) Data

1. Laden Sie die Datei `EVChargeData.csv` aus der zip-Datei `daten.zip` von energie.labs.fhv.at/~kr/effsys-eu/unterlagen.html herunter. Datenquelle: The project *My Electric Avenue* monitored the usage of over 200 EVs. cf. *My Electric Avenue Published Technical Data Overview.pdf* in the download: *My Electric*

- Avenue Technical Data.tar.xz at eatechnology.com. Alternative EV-Daten gibt's z. B. unter [Electric Chargepoint Analysis 2017](#).
- Laden Sie die ersten 5000 Datenzeilen von `EVChargeData.csv` in ein pandas dataframe, und parsen Sie die Zeitangaben als Timestamps.
 - Ändern Sie die Spaltennamen zu `'ID'`, `'start'`, `'stop'`, `'SOC_start'`, `'SOC_stop'`.
 - Fügen Sie zwei Spalten mit Namen `duration` und `SOC_charged` an, die die Ladedauer in Stunden bzw. die Lademenge in den SOC-Einheiten von 0 - 12 angeben.
 - Erstellen Sie ein Histogramm der Häufigkeiten der Lademengen.
 - Erstellen Sie einen Scatterplot für die zwei Features `duration` und `SOC_charged`:
 - Erstellen Sie zusätzliche Spalten für
 - Stunde des Ladebeginns
 - Stunde des Ladeendes
 - Dauer des Ladens in ganzen Stunden
 - Speichern Sie in einem pandas dataframe die Wahrscheinlichkeiten, in einer bestimmten Dauer (in ganzen Stunden) eine bestimmte SOC-Menge zu laden, und Stellen Sie diese Daten grafisch dar.
 - Bestimmen Sie die bedingten Wahrscheinlichkeiten $P(\text{SOC_charged}=x|\text{duration}=3)$, und stellen Sie diese als bar plot über x dar.

2.11. Szenariogeneration für Simulationen

Wenn man eine unsichere (Mess)Größe durch eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}$ modelliert hat, kann man durch Ziehen von Ergebnissen (Szenarien, Stichproben) $\omega \in \Omega$ die zugehörigen Szenariowerte $X(\omega)$ generieren. Diese Simulationsmethode wird Szenariogeneration, [Monte Carlo Simulation](#) oder "random number generation" genannt. Da die Verteilung bzw. die Verteilungsdichte (= Wahrscheinlichkeitsdichte) der Messgröße entscheidend ist, spricht man auch von der Szenariogeneration aus bestimmten Verteilungen.

Zu Beginn initialisieren wir einen Zufallszahlengenerator, Englisch *random number generator*:

```
# construct a random number generator (rng):
from numpy.random import default_rng
rng = default_rng()
```

Wir ziehen Samples/Szenarien. Hier sind es 10 zufällige ganze Zahlen, die gleichmäßig zwischen 0 (inklusive) und 5 (exklusive) verteilt sind.

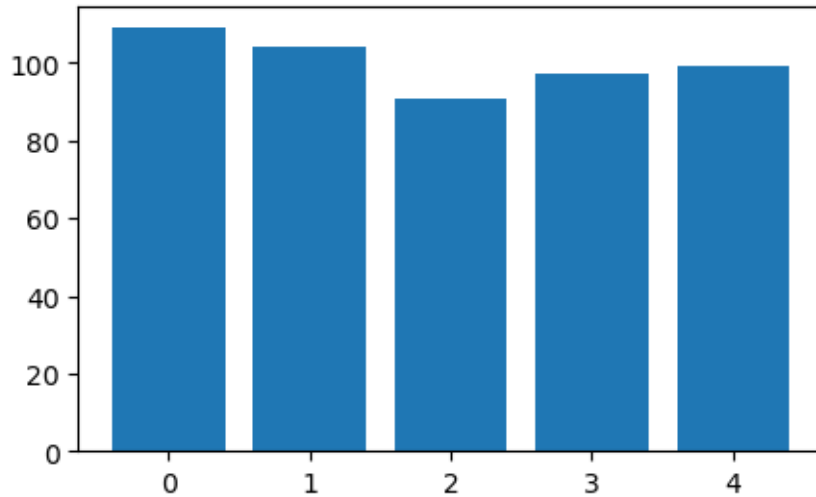
```
# drawing random integers that are uniformly distributed:
low = 0 # lowest integer
high = 5 # highest integer
rng.integers(low, high, size=20)
```

```
array([4, 1, 0, 0, 0, 0, 4, 3, 2, 1, 1, 2, 4, 1, 1, 1, 4, 0, 2, 3])
```

Histogramm/Verteilung von Samples/Szenarien:

```
# make histogram for given samples (=scenarios):
samples = rng.integers(low, high, size=500)

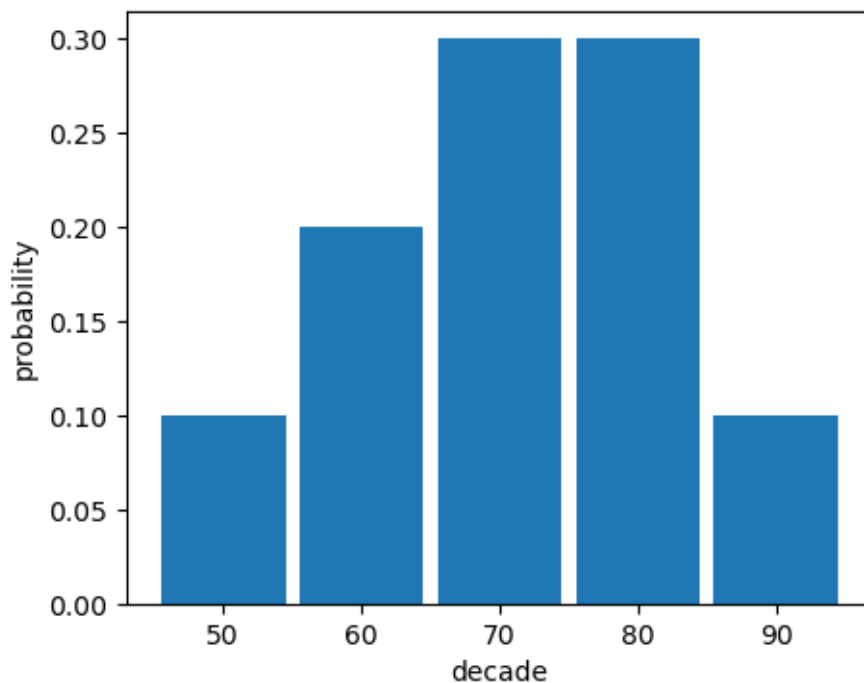
plt.figure(figsize=(5, 3))
plt.hist(samples, bins=range(low, high + 1), rwidth=0.8, align='left');
```



Samples/Szenarien aus einer Verteilung/Histogramm:

```
# Example: random variable = life expectancy in decades
x = np.array([50, 60, 70, 80, 90]) # decades of death
p = np.array([0.1, 0.2, 0.3, 0.3, 0.1]) # probability
```

```
# histogram of random variable:
plt.figure(figsize=(5, 4))
plt.bar(x, p, width=9)
plt.xticks(x)
plt.xlabel('decade')
plt.ylabel('probability');
```



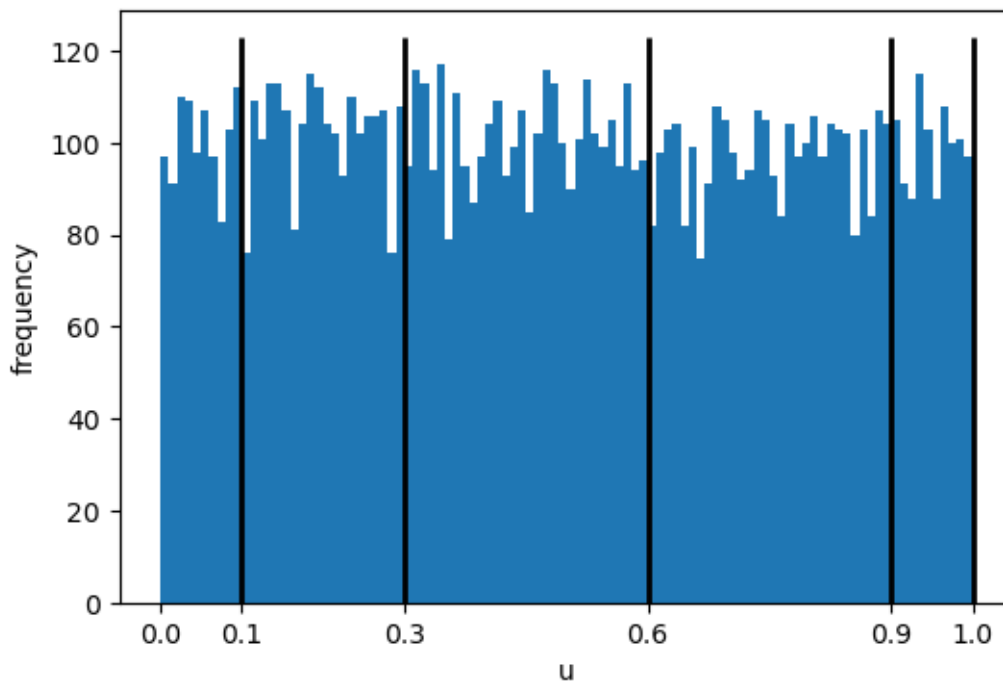
```
# cumulative sum of probabilities:
P = p.cumsum()
P
```

```
array([0.1, 0.3, 0.6, 0.9, 1. ])
```

Wir generieren zuerst n kontinuierliche Zufallszahlen zwischen 0 und 1:

```
n = 10000
u = rng.uniform(size=n)

plt.figure(figsize=(6, 4))
plt.hist(u, bins=100)
plt.xlabel('u')
plt.ylabel('frequency')
plt.xticks(np.concatenate((np.array([0]), P)))
plt.xlim(-0.05, 1.05)
y_min, y_max = plt.ylim()
for cum_prob in P:
    plt.vlines(cum_prob, y_min, y_max, colors="k", linewidth=2)
```



- Jene Zufallszahlen, die zwischen 0.0 und 0.1 liegen machen ca. $0.1 = 10\%$ aller Zufallszahlen aus. Diese ordnen wir $x = 50$ zu.
- Jene Zufallszahlen, die zwischen 0.1 und 0.3 liegen machen ca. $0.2 = 20\%$ aller Zufallszahlen aus. Diese ordnen wir $x = 60$ zu.
- Jene Zufallszahlen, die zwischen 0.3 und 0.6 liegen machen ca. $0.3 = 30\%$ aller Zufallszahlen aus. Diese ordnen wir $x = 70$ zu.
- Jene Zufallszahlen, die zwischen 0.6 und 0.9 liegen machen ca. $0.3 = 30\%$ aller Zufallszahlen aus. Diese ordnen wir $x = 80$ zu.
- Jene Zufallszahlen, die zwischen 0.9 und 1.0 liegen machen ca. $0.1 = 10\%$ aller Zufallszahlen aus. Diese ordnen wir $x = 90$ zu.

```
x_samples = x[np.searchsorted(P, u)]

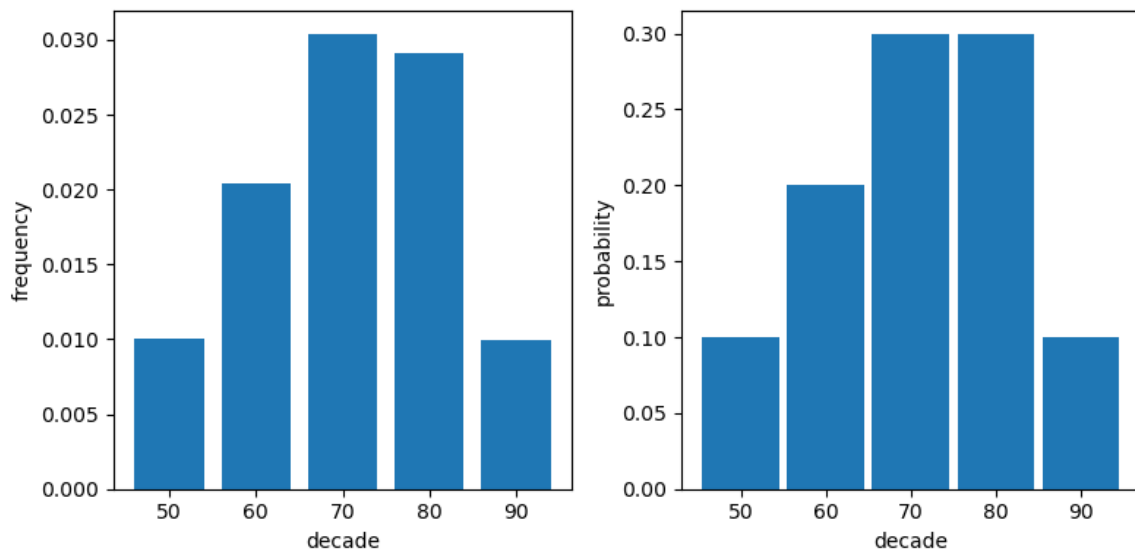
# print first 10 u and x samples:
for i in range(10):
    print(f"u[{i}] = {u[i]:.2f}, x_samples[{i}] = {x_samples[i]}")
```

```
u[0] = 0.33, x_samples[0] = 70
u[1] = 0.91, x_samples[1] = 90
u[2] = 0.75, x_samples[2] = 80
u[3] = 0.04, x_samples[3] = 50
u[4] = 0.76, x_samples[4] = 80
u[5] = 0.04, x_samples[5] = 50
u[6] = 0.57, x_samples[6] = 70
u[7] = 0.31, x_samples[7] = 70
u[8] = 0.32, x_samples[8] = 70
u[9] = 0.70, x_samples[9] = 80
```

```
plt.figure(figsize=(8, 4))
plt.subplot(1, 2, 1)
plt.hist(x_samples, bins=[45, 55, 65, 75, 85, 95, 95],
         rwidth=0.8, density=True)
plt.xlabel('decade')
plt.ylabel('frequency')

# compare with the histogram of random variable:
plt.subplot(1, 2, 2)
plt.bar(x, p, width=9)
plt.xticks(x)
plt.xlabel('decade')
plt.ylabel('probability')

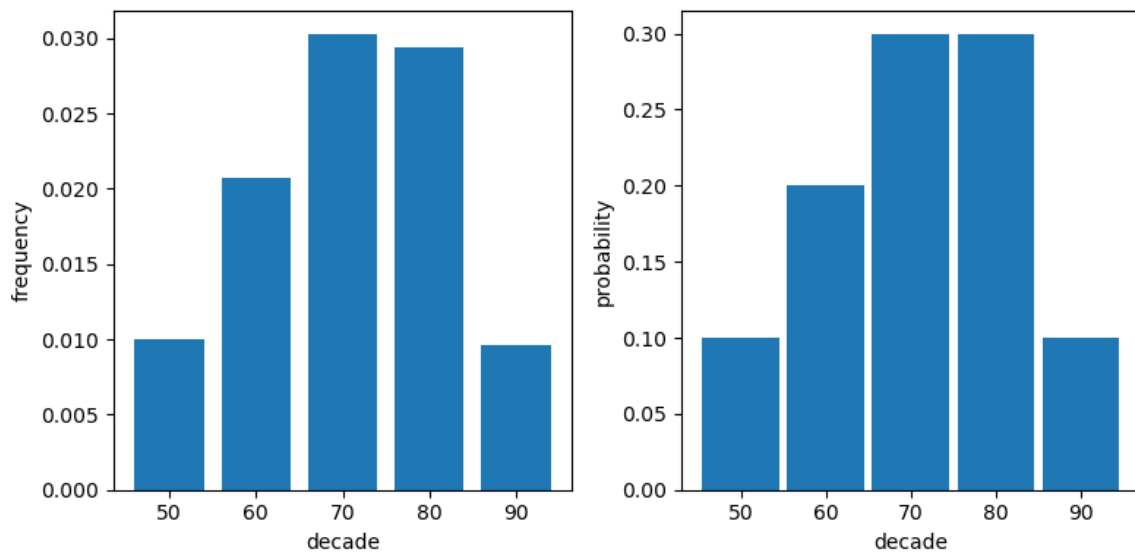
plt.tight_layout()
```



Schnelle Variante mit numpy-Funktion choice:


```
x_samples = rng.choice(x, p=p, size=n)
```

```
plt.figure(figsize=(8, 4))  
plt.subplot(1, 2, 1)  
plt.hist(x_samples, bins=[45, 55, 65, 75, 85, 95],  
         rwidth=0.8, density=True)  
plt.xlabel('decade')  
plt.ylabel('frequency')  
  
# compare with the histogram of random variable:  
plt.subplot(1, 2, 2)  
plt.bar(x, p, width=9)  
plt.xticks(x)  
plt.xlabel('decade')  
plt.ylabel('probability')  
  
plt.tight_layout()
```



Teil II.

Aufgaben

3. Aufgaben 1

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.

3.1. XXX (X Punkte)

4. Aufgaben 2

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.

4.1. XXX (X Punkte)

5. Aufgaben 3

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.

6. Aufgaben 4

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.

7. Aufgaben 5

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.

8. Aufgaben 6

Die Aufgaben müssen in [LIAS](#) als lauffähige und dokumentierte Jupyter Notebooks inkl. Datenfiles abgegeben werden. Handschriftliche Modellierungen und Grafiken können als eingescannte PDFs abgegeben werden.