Serie 10

Aufgabe 10.1

Der Fachartikel *Compression of Single-Wall Corrugated Containers Using Fixed and Floating Test Platens* (J. Testing and Evaluation, 1992: 318-320) beschreibt ein Experiment, in dem verschiedene Typen von Container-Hüllen in Bezug auf Druckfestigkeit (lb) verglichen wurden.

Тур	Druckfestigkeit					
1	655.5	788.3	734.3	721.4	679.1	699.4
2	789.2	772.5	786.9	686.1	732.1	774.8
3	737.1	639.0	696.3	671.7	717.2	727.1
4	535.1	628.7	542.4	559.0	586.9	520.0

a) Geben Sie die Daten selber in **Python** ein, und stellen Sie sie mit Stripcharts und Boxplots dar.

Python-Hinweise: Die Daten werden in ein Dataframe mit zwei Spalten eingelesen: eine Spalte mit Druckfestigkeitsangaben und eine Spalte mit Hüllentyp:

```
from pandas import DataFrame
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import scipy.stats as st
df=DataFrame({
  "Typ": np.repeat(["T1", "T2", "T3", "T4"], [6, 6, 6, 6]),
  "Druckfestigkeit" : [655.5, 788.3, 734.3, ..]
})
sns.stripplot(x="Typ", y="Druckfestigkeit", data=df)
plt.xlabel("Typ")
plt.ylabel("Druckfestigkeit")
plt.show()
sns.boxplot(x="Typ", y="Druckfestigkeit", data=df)
plt.xlabel("Typ")
plt.ylabel("Druckfestigkeit")
plt.show()
```

- b) Wie lautet ein Gruppenmittelmodell passend zum Datensatz und zur Fragestellung? Schätzen Sie die Parameter Ihres Modelles.
- c) Besteht ein Unterschied zwischen den Hüllentypen? Führen Sie einen statistischen Hypothesentest auf dem 5 % Niveau durch.

Aufgabe 10.2

24 Tiere werden zufällig zu 4 unterschiedlichen Ernährungsdiäten zugeordnet, um den Effekt auf die Blutkoagulationszeit zu untersuchen.

```
Behandlung Koagulationszeit

A 62 60 63 59

B 63 67 71 64 65 66

C 68 66 71 67 68 68

D 56 62 60 61 63 64 63 59
```

a) Geben Sie die Daten selber in **Python** ein, und stellen Sie sie mit Stripcharts und Boxplots dar.

Python-Hinweise: Die Daten werden in ein Dataframe mit zwei Spalten eingelesen: eine Spalte mit Druckfestigkeitsangaben und eine Spalte mit Hüllentyp:

```
from pandas import DataFrame
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import scipy.stats as st
df=DataFrame({
  "Behandlung": np.repeat(["A", "B", "C", "D"], [4, 6, 6, 8]),
  "Koagulationszeit" : [62, 60, 63, ..]
})
sns.stripplot(x="Behandlung", y="Koagulationszeit", data=df)
plt.xlabel("Behandlung")
plt.ylabel("Koagulationszeit")
plt.show()
sns.boxplot(x="Behandlung", y="Koaqulationszeit", data=df)
plt.xlabel("Behandlung")
plt.ylabel("Koagulationszeit")
plt.show()
```

- b) Berechnen Sie den globalen Mittelwert (grand mean) und die Gruppenmittelwerte mit Hilfe eines Taschenrechners.
- c) Berechnen Sie mit Hilfe eines Taschenrechners die empirischen Gruppenvarianzen

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} r_{ij}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - y_{i.})^2$$

- d) Berechnen Sie die pooled Varianz S_{Pool}^2 , resp. MS_E .
- e) Berechnen Sie MS_G und vergleichen Sie den Wert mit MS_E.
- f) Konstruieren Sie eine ANOVA Tabelle mit Hilfe von Python.
- g) Besteht ein signifikanter Unterschied zwischen den Behandlungsarten in Bezug auf die Koagulationszeit? Führen Sie einen statistischen Hypothesentest auf dem 5 % Niveau durch.

Kurzlösungen vereinzelter Aufgaben

A 10.1:

c) P-Wert 5.5*e* − 7

A 10.2:

- b) $\mu = 64$ $\mu_A = 61$ $\mu_B = 66$ $\mu_C = 68$ $\mu_D = 61$
- c) $s_A^2 = 3.333$ $s_B^2 = 8$ $s_C^2 = 2.8$ $s_D^2 = 6.85$
- d) $MS_E = 5.6$
- e) $MS_G = 76$