



Übung zu Digitalisierung und Statistik

Abgabe & Bearbeitung: Mittwoch, 15. Jänner

Aufgabe 12.1 (5 Punkte)

Gegeben sei die Excel-Datei „Messdaten-Mittwoch.xlsx“. Laden Sie das erste Datenblatt mittels Pandas in Python ein. Sie werden feststellen, dass das die selben Daten sind wie in der ersten Aufgabe der letzten Woche. Wir wollen die Regressionsanalyse jetzt vervollständigen. Berechnen Sie erneut für die beiden ersten Spalten die Parameter m , b einer linearen Regression sowie das Bestimmungsmaß R^2 , indem Sie Ihren Code von letzter Woche wiederverwenden. Plotten Sie in **einem** Plot die Punkte aus dem Datenblatt mit der `scatter` Funktion und geben Sie den Daten das Label „Gemessene Punkte“. Plotten Sie mit der `plot` Funktion den Funktionsgraphen der linearen Funktion mit den geschätzten Parametern. Geben Sie dem Funktionsgraphen das Label „ $y = \{m\}x + \{b\}$, $R^2 = \{r2\}$ “. Fügen Sie natürlich in die Stellen $\{m\}$, $\{b\}$ und $\{r2\}$ die richtigen Zahlenwerte ein.

Achtung: Wenn Sie zwei (oder mehr) unabhängige Plots in einem Skript nacheinander erstellen möchten, müssen Sie Matplotlib nach jedem Plot resettet, ansonsten werden die Daten des zweiten Plots mit dem ersten überlagert. Das erreichen Sie mit diesem Befehl:

```
plt.close()
```

Aufgabe 12.2 (4 Punkte)

Drei verschiedene Kopfschmerztabletten wurden auf ihre Wirkdauer (in Minuten) an verschiedenen Personen geprüft. Erstellen Sie **einen** Plot, in dem die Wirkdauern der Tabletten als Histogramm geplottet sind. Nutzen Sie dazu die Daten aus der angehängten Datei „Messdaten-Mittwoch.xlsx“ aus dem Arbeitsblatt „Wirkdauer“. Die Histogramm-Balken sollen dabei halbtransparent sein. Dies kann bei der `hist` Funktion mit dem optionalen Parameter `alpha = 0.5` erreicht werden. Lesen Sie die Daten mit Pandas ein und plotten Sie sie mittels Matplotlib.

Aufgabe 12.3 (1 Punkt)

Erweitern Sie den Plot von Aufgabe 12.2 indem Sie über alle drei Histogramme die Dichtefunktion der Normalverteilung drüberlegen.