

Übungsblatt: Poisson- und Binomialverteilung

DSA Rossleben – Kurs 5.1

Aufgabe 13.5

In dieser Aufgabe erkunden wir die Poissonverteilung mit Python. Du darfst in dieser Aufgabe (und auch allen folgenden Aufgaben) natürlich gerne auch das Internet zu Rate ziehen, wenn du noch Fragen hast, wie man in Python programmiert. (Simon und Carlo machen das auch so.)

- Erzeuge mit `numpy.random.poisson( $\lambda=5$ , size=10000)` einen Datensatz von 10 000 Poisson-Zufallszahlen. Plote das Histogramm (normiert, `density=True`) und vergleiche es mit der theoretischen Poisson-Verteilung $P(k; \lambda)$ aus dem Text.
- Berechne für diesen Datensatz den Stichprobenmittelwert $\bar{k}$ und die Stichprobenstandardabweichung s . Vergleiche mit den theoretischen Werten $\langle k \rangle = \lambda$ und $\sigma_k = \sqrt{\lambda}$.

Aufgabe 13.6

Wir haben eine exotische Münze, die mit Wahrscheinlichkeit p Kopf zeigt. Die Münze wird 14-mal geworfen; dabei werden 10-mal Kopf und 4-mal Zahl beobachtet (Daten D). Wir interessieren uns für die Wahrscheinlichkeit, dass die nächsten beiden Würfe Kopf zeigen.

- Beschreibe das statistische Modell für dieses Experiment.
- Der sogenannte **Maximum-Likelihood-Schätzer** aus der frequentistischen Statistik schätzt die unbekannte Wahrscheinlichkeit p durch

$$\hat{p} = \frac{\text{Anzahl Kopf}}{\text{Anzahl Würfe}}. \quad (1)$$

Berechne den Schätzwert $\hat{p}$ und bestimme damit die Wahrscheinlichkeit, dass die nächsten beiden Würfe Kopf zeigen.

- Schreibe in Python eine Funktion, die für ein gegebenes $p = 5/7$ aufeinander folgende 14 Würfe simuliert. Werte diese Funktion 10 000-mal aus und mache eine Liste der Male, in denen genau 10-mal Kopf und 4-mal Zahl beobachtet wurde. Welche frequentistische Wahrscheinlichkeit hat also dieses Ergebnis (10-mal Kopf und 4-mal Zahl in 14 Würfeln)?

Hinweis: Die Funktion `np.random.binomial(n, p)` zieht dir direkt eine solche Stichprobe. Google gerne mal, wie man diese importiert und wofür n dabei steht.

- Bei n unabhängigen Versuchen mit Erfolgswahrscheinlichkeit p folgt die Anzahl der Erfolge k der Verteilung

$$B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}. \quad (2)$$

Schreibe in Python eine Funktion, die die Wahrscheinlichkeit für k -mal Kopf bei n Würfeln mit vorgegebenem p berechnet. Berechne damit für

$$p = 0.2, 0.4, 0.6, 0.8, 0.9$$

jeweils die Wahrscheinlichkeit, bei 14 Würfeln genau 10-mal Kopf und 4-mal Zahl zu beobachten.

Hinweis: Hier ist `np.random.binomial` *nicht* das richtige Werkzeug – das liefert dir nur eine zufällige Stichprobe, aber keine Wahrscheinlichkeit. Berechne stattdessen $n!$ selbst (oder mit `math.factorial`) und setze damit den Binomialkoeffizienten und $B(k; n, p)$ direkt aus der obigen Formel zusammen.

- e) Plote die Ergebnisse der Aufgaben c) und d) gemeinsam und vergleiche sie miteinander. Da c) bisher nur eine einzige frequentistische Wahrscheinlichkeit bei $p = \hat{p}$ liefert, musst du die Simulation aus c) dafür zunächst erweitern: Wiederhole sie für jeden der fünf Werte $p = 0.2, 0.4, 0.6, 0.8, 0.9$ aus Teil d). Trage anschließend sowohl die theoretischen (Teil d) als auch die simulierten (Teil c) Wahrscheinlichkeiten für $P(10\text{-mal Kopf})$ über p auf.
- f) (*Bonus*) In der Bayesschen Statistik wird p als zufällige Größe modelliert. Wir nehmen zunächst an, dass alle Werte zwischen 0 und 1 gleich wahrscheinlich sind (*a-priori*-Verteilung). Versuche, die *a-posteriori*-Verteilung (numerisch oder auf dem Papier) zu berechnen. Wie groß ist die Wahrscheinlichkeit, dass die nächsten beiden Würfe Kopf zeigen, wenn du diese Verteilung zur Berechnung verwendest? Unterscheidet sich das von dem frequentistischen Ergebnis? Basierend auf deiner Analyse: Solltest du Geld darauf verwetten, dass die nächsten beiden Würfe Kopf zeigen?

Hinweis: Erwartungswerte werden wie folgt berechnet:

$$\langle f(p) \rangle = \int_0^1 f(p) P(p | D) dp. \quad (3)$$

Wenn du noch nicht gelernt hast, wie man das relevante Integral löst, recherchiere gerne im Internet, wie das numerisch in `Python` geht.