ETH Zürich HS 2021
Neur. Syst. Comp., D-MATL, D-MAVT
Patrick Cheridito

Koordinatorin
Hanna Wutte

# Stochastik

## Serie 9

**1.** Auf einem Weihnachtsmarkt findet Laura einen Stand mit einer Maschine, die das mehrfache Werfen einer Münze durchführt und das Ergebnis ausliest. So können in kurzer Zeit viele Münzwürfe durchgeführt werden. Beantworte für jede der nachfolgenden Teilaufgaben:
Für welche der je zwei Optionen sollte sich Laura entscheiden: (1), (2), oder sind beide Optionen etwa gleich gut? Bedenke, dass sie keinen Taschenrechner dabei hat, also die Gewinnwahrscheinlichkeiten nicht genau ausrechnen kann.
**Tipp:** Überlege wie sich die Standardabweichung für $\bar{X}_n$ und $S_n$ mit $n$ ändern.

**a)** (1) Es wird 100 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze mehr als 52% Kopf zeigt.

(2) Es wird 1000 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze mehr als 52% Kopf zeigt.

**b)** (1) Es wird 100 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze weniger als 55 mal Kopf zeigt.

(2) Es wird 200 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze weniger als 105 mal Kopf zeigt.

**c)** (1) Es wird 100 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze weniger als 40 mal Kopf zeigt.

(2) Es wird 400 mal eine Münze geworfen und Laura bekommt 10 CHF, wenn die Münze weniger als 180 mal Kopf zeigt.

**2.** Für eine Studie werden über eine Telefonumfrage zufällig ausgewählte Personen über sensitive Bereiche in ihrem Leben, wie zum Beispiel Drogenkonsum befragt. Bei einer direkten Frage, wird solch eine Frage möglicherweise mit "nein" beantwortet, obwohl die Person Drogen konsumiert hat. Oft wird einfach die sozial besser akzeptierte Anwort gewählt. Um dies zu vermeiden kann man statt direkter Nachfrage wie folgt vorgehen:
Bitte die Person eine Münze zu werfen und sich das Resultat zu merken - dies aber für sich zu behalten. Die Person soll nun, wenn die Münze Kopf gezeigt hat mit "ja" antworten, und wenn die Münze Zahl gezeigt hat, die Frage beantworten.

**Bitte wenden!**

**a)** Warum ist es wahrscheinlicher, dass Leute auf diesem Weg die Wahrheit sagen?

**b)** Nehmen wir eine zufällige Auswahl von 100 Personen aus der Bevölkerung und befragen diese auf dem oben vorgeschlagenen Weg nach dem Drogenkonsum im letzten Jahr. Angenommen 65 Personen antworten mit "ja". Berechne den Momentenschätzer für den Anteil der Bevölkerung, der im vergangenen Jahr Drogen konsumiert hat.

**3.** Um genauere Prognosen bezüglich der Flugzeiten ihrer Flugzeuge machen zu können, möchte die Fluggesellschaft SwitzAir zur Modellierung der absoluten Windgeschwindigkeiten neuerdings eine sogenannte Rayleigh-Verteilung mit Dichte

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) & x \geq 0 \end{cases}$$

und kumulativer Verteilungsfunktion

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right) & x \geq 0 \end{cases}$$

benutzen, wobei $\sigma > 0$ ein unbekannter Parameter ist. Die Forschungsabteilung beauftragt uns, den unbekannten Parameter $\sigma$ der Verteilung anhand von $n$ beobachteten Datenpunkten zu schätzen.

**Tipp:** Für eine Rayleigh($\sigma$)-verteilte Zufallsvariable $X$ gilt $\mathbb{E}[X] = \sqrt{\frac{\pi}{2}}\sigma$ und $\mathbb{V}(X) = \frac{4-\pi}{2}\sigma^2$.

**a)** Bestimme die Likelihood- und die log-Likelihoodfunktion basierend auf den $n$ unabhängigen, identisch verteilten Beobachtungen $x_1, \ldots, x_n$ einer Zufallsvariable mit obiger Dichte. Berechne daraus den Maximum-Likelihood Schätzer für den unbekannten Parameter $\sigma$.

Da die SwitzAir selbst über keine Messinstrumente zur Messung der absoluten Windgeschwindigkeiten verfügt, beauftragt sie ein meteorologisches Institut, 100 Windgeschwindigkeitsmessungen durchzuführen. Das meteorologische Institut liefert folgende Kennzahlen: Mittelwert $\bar{x} = 2.01$, Standardabweichung $s_x = 0.51$, Median $m = 3.12$ und Quartilsdifferenz $q_{0.75} - q_{0.25} = 0.89$.

**Siehe nächstes Blatt!**

**b)** Berechne den Momentenschätzer basierend auf dem ersten Moment. Welchen Wert nimmt der Schätzer für die gesammelten Daten des meteorologischen Instituts an?

**4.** Wir nehmen an, dass der Bleigehalt $X$ in einer Bodenprobe normalverteilt ist,

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

**a)** Die Standardabweichung ist bekannt ($\sigma = 10$ ppb (parts per billion)) und man interessiert sich für den Erwartungswert $\mu$. Man misst den Bleigehalt in 10 Bodenproben und erhält einen Mittelwert von 31 ppb. Gib ein 99% Vertrauensintervall für $\mu$ an.

**Tipp**: Die kumulative Verteilungsfunktion der Standardnormalverteilung erfüllt
$\Phi(2.58) = 0.995$.

**b)** Wieviel Beobachtungen sind nötig, um die Breite des Vertrauensintervalles auf die Hälfte zu reduzieren?

Wieviele (unabhängige) Bestimmungen des Bleigehaltes müssen geplant werden, wenn der Bleigehalt mit einer Stichprobe "auf 1 ppb genau" bestimmt werden soll, d.h. wenn die Breite des 99 % Konfidenzintervalls nicht grösser als 1 ppb sein soll?