

## Aufgabenblatt 2 – Lösungen

### Aufgabe 1 (Kennzahlen)

- a) Die Schlusskurse von Bitcoin und Ethereum aus 2022 haben eine Kovarianz von 7718820. Kann daraus ein starker, positiver Zusammenhang geschlossen werden?

Eine positive Kovarianz zeigt einen positiven Zusammenhang an. Sie ist aber nicht normiert, daher bleibt die relative Stärke des Zusammenhanges unbekannt.

- b) Finden Sie ein Beispiel, dass im Allgemeinen  $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$  für zwei Zufallsvariablen  $X$  und  $Y$  gilt.

Sei  $X = Y \sim N(0, 1)$ . Dann ist  $\text{Var}(X) = \text{Var}(Y) = 1$ , aber  $\text{Var}(X + Y) = 4$ .

- c) Finden Sie zwei Zufallsvariablen, die unkorreliert, aber nicht unabhängig sind.

Wähle  $X, Y$  unabhängig mit  $\Pr(X = 0) = \Pr(X = 1) = 0.5$ ,  $\Pr(Y = -1) = \Pr(Y = 1) = 0.5$  und definiere  $Z := XY$ . Dann ist  $\text{Cov}(Z, X) = E(Z(X - 0.5)) = E(X^2 - 0.5X)E(Y) = 0$  aber  $\Pr(Z = 1 \mid X = 0) = 0 \neq 0.5 = \Pr(Z = 1 \mid X = 1)$ .

- d) Finden Sie zwei reelle Vektoren  $x$  und  $y$ , jeweils der Länge 10, sodass die empirische Korrelation  $\widehat{\text{Cor}}(x, y)$  exakt  $-1$  bzw.  $0$  bzw.  $+1$  beträgt.

Für  $x = 1, \dots, 10$  und  $y = x$  ist  $\widehat{\text{Cor}}(x, y) = +1$ . Für  $y = -x$  ist  $\widehat{\text{Cor}}(x, y) = -1$ . Für  $y$  mit  $y_i = (x - \bar{x})^2$ ,  $i = 1, \dots, 10$  ist  $\widehat{\text{Cor}}(x, y) = 0$ .

- e) Es seien  $X, Y$  zwei unabhängige Zufallsvariablen, die mit gleicher Wahrscheinlichkeit die Werte in  $\{1, 2, 3\}$  annehmen. Berechnen Sie  $E(1 + 4X + 2Y \mid X = 2)$ .

$E(1 + 4X + 2Y \mid X = 2) = 9 + 2E(Y) = 13$

### Aufgabe 2 (Schließende Statistik)

Ein Losverkäufer behauptet, dass mindestens 20% seiner Lose Gewinne seien. Die Käufer aber vermuten, dass der Anteil geringer ist. Es werden  $n = 100$  Lose überprüft.

- a) Führen Sie einen statistischen Test zum Signifikanzniveau  $\alpha = 5\%$  zur Streitschlichtung durch, wobei Sie die Aussage des Losverkäufers als Nullhypothese wählen.

Der Anteil an Gewinnlosen sei binomialverteilt zu  $n = 100$  und  $p = 20\%$ . Die Wahrscheinlichkeit, höchstens 14 Erfolge zu haben, beträgt somit 8%, und höchstens 13 Erfolge zu haben, 4.7%. Zu  $\alpha = 5\%$  geben wir also dem Losverkäufer recht, falls in der Stichprobe mindestens 14 Gewinne sind, ansonsten den Käufern. Siehe <https://www.youtube.com/watch?v=MBf9Iin6bpg> für mehr Details.

- b) Angenommen, der wahre Anteil an Gewinnlosen beträgt nur 10%. Berechnen Sie die Wahrscheinlichkeit, mit der Ihr Test dem Losverkäufer fälschlicherweise recht gibt.

Das ist die Wahrscheinlichkeit, unter der Binomialverteilung mit  $n = 100$  und  $p = 10\%$  mindestens 14 Erfolge zu beobachten. Sie beträgt 12.4%.

- c) Sie möchten die in b) berechnete Wahrscheinlichkeit auf unter 5% reduzieren. Auf welche Werte müssen Sie dafür entweder  $\alpha$  oder  $n$  verändern?

Entweder  $\alpha$  auf mindestens 13% oder  $n$  auf mindestens 135 erhöhen.

### Aufgabe 3 (Vorhersagen)

- a) Überlegen Sie, wie präzise Folgendes vorhersagbar ist: der morgige Bitcoinpreis, die Lebensdauer einer Batterie, ein Münzwurf, die nächste Sonnenfinsternis, Ihr späteres Gehalt, das Wetter am nächsten Wochenende.

Der morgige Bitcoinpreis ist nicht bis kaum vorhersagbar. Die Lebensdauer einer Batterie ist mit einer gewissen Unsicherheit vorhersagbar. Das Ergebnis eines Münzwurfs ist nicht vorhersagbar. Der Zeitpunkt der nächsten Sonnenfinsternis ist exakt vorhersagbar. Mein späteres Gehalt ist mit einer gewissen Unsicherheit vorhersagbar, ebenso das Wetter am nächsten Wochenende.

- b) Beschreiben Sie drei Alltagssituationen, in denen Sie Vorhersagen treffen. Warum glauben Sie, dass diese Situation prognostizierbar sind? Welche Faktoren führen zu Prognosefehlern? Welchen Wert haben gute im Vergleich zu schlechten Vorhersagen für Sie?

Zum Beispiel das benötigte Geld für den nächsten Supermarkteinkauf: meine Einkaufsliste hilft, spontane Einkäufe oder Preisänderungen führen zu Fehlern, eine gute Vorhersage hilft bei der Haushaltsplanung, eine Unterschätzung möglicherweise zu Engpässen am Monatsende.

- c) Erklären Sie anhand eines Beispiels, dass falsche Vorhersagen unterschiedliche Konsequenzen nach sich ziehen können.

Abflugzeit: zu früh eingeschätzt ist meistens kostengünstiger als zu spät.

- d) Welche Charakteristiken weisen Zeitreihen auf, die leichter beziehungsweise schwieriger vorhersagbar sind?

Bei der Vorhersage hilft eine Struktur in der Zeitreihe, wie Trend und Saisonalität, oder ein starker Zusammenhang mit einer anderen Variable, die beobachtbar ist.