# Stochastik
# (BSc D-MAVT / BSc D-MATH / BSc D-MATL)

**Schreiben Sie für Aufgabe 2-4 stets alle Zwischenschritte und -rechnungen sowie Begründungen auf. Aufgabe 1 ist eine Multiple Choice Aufgabe (keine Begründungen notwendig).**

**Die für die Aufgaben benötigte Tabelle der Normalverteilung wurde mit der Prüfung ausgeteilt.**

**1. (10 Punkte)**

Bei den folgenden 10 Fragen ist jeweils genau eine Antwort richtig. Es gibt pro richtig beantwortete Frage 1 Punkt und pro falsche Antwort 1/2 Punkt Abzug. Minimal erhält man für die gesamte Aufgabe 0 Punkte. **Bitte benützen Sie das beiliegende Antwortblatt.**

**a)** Eine Münze wird zweimal geworfen. Für $i = 1, 2$ sei $K_i$ das Ereignis, dass beim $i$-ten Wurf Kopf erscheint. Dann sind die Ereignisse $K_1$ und $K_2$

1. disjunkt, aber nicht unabhängig.
2. unabhängig, aber nicht disjunkt.
3. disjunkt und unabhängig.

**b)** Wann gilt das Additionsgesetz $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$?

1. Falls $A$ und $B$ disjunkt sind.
2. Falls $A$ und $B$ unabhängig sind.
3. Das Additionsgesetz gilt ohne weitere Voraussetzungen.

**c)** Seien $A$ und $B$ Ereignisse mit $\mathbb{P}(A \cap B) = \frac{1}{4}$ und $\mathbb{P}(B) = \frac{1}{2}$. Daraus kann man schliessen:

1. $\mathbb{P}(A) \in [0, \frac{1}{4}]$.
2. $\mathbb{P}(A) \in [\frac{1}{4}, \frac{3}{4}]$.
3. $\mathbb{P}(A) \in [\frac{3}{8}, 1]$.

**Bitte wenden!**

**d)** Seien $T, H$ zwei unkorrelierte Zufallsvariablen mit $\mathbb{E}(T) = \mathbb{E}(H) = 1$ und $\mathrm{Var}(T) = \mathrm{Var}(H) = 1$. Dann gilt:

1. $\mathbb{E}\left((T+H)^2\right) = 2.$
2. $\mathbb{E}(TH) = 2.$
3. $\mathbb{E}\left((T-H)^2\right) = 2.$

**e)** Durch welche der folgenden Verteilungen lässt sich die Anzahl der Erdbeben in der Schweiz innerhalb eines Jahres am besten modellieren?

1. Poissonverteilung.
2. Binomialverteilung.
3. Normalverteilung.

**f)** Eine faire Münze wird mehrfach geworfen. Wie gross ist die Wahrscheinlichkeit, dass beim 41. Wurf Zahl erscheint, gegeben, dass bei den ersten 40 Würfen jeweils Zahl erschienen ist?

1. $(\frac{1}{2})^{41}$.
2. $\frac{1}{41}$.
3. $\frac{1}{2}$.

**g)** Wir betrachten eine Stichprobe $x_1, \ldots, x_n$ bestehend aus positiven, reellen Zahlen. Welche Kennzahl lässt sich aus einem Histogramm dieser Stichprobe über den Klassen $(0, 1.5], (1.5, 2.5], (2.5, 3.5], \ldots$ bestimmen?

1. Der empirische Mittelwert.
2. Der empirische Median.
3. Im Allgemeinen lässt sich weder der empirische Mittelwert noch der empirische Median aus dem Histogramm ablesen.

**h)** Welche der folgenden Grössen eines Tests hängt von der gegebenen Stichprobe $x_1, \ldots, x_n$ ab?

1. Das Signifikanzniveau.
2. Der p-Wert.
3. Der Verwerfungsbereich.

**i)** Seien $a < b$ reelle Zahlen. Wir nehmen an $I = [a, b]$ sei das 95%-Vertrauensintervall für den Mittelwert von normalverteilten Daten mit Standardabweichung 0.05, das mit Hilfe eines (zweiseitigen) $Z$-Tests aus 100 Datenpunkten ermittelt wurde. Wieviele Datenpunkte bräuchte man, um die Länge $b - a$ des Vertrauensintervalls zu halbieren?

1. 200 Datenpunkte.
2. 400 Datenpunkte.
3. Die Länge des Vertrauensintervalls hängt von den Beobachtungen ab und kann daher nicht allein durch die Anzahl der Datenpunkte bestimmt werden.

**Siehe nächstes Blatt!**

**j)** In einer Studie wird untersucht, ob eine Erhöhung der Raumtemperatur von 22 °C auf 28 °C die geistige Leistungsfähigkeit des Menschen beeinflusst. Eine relative geistige Leistungsfähigkeit von 100% bedeutet, dass man bei beiden Temperaturen gleich leistungsfähig ist. Dies ist die Nullhypothese. Eine relative geistige Leistungsfähigkeit von mehr als 105% oder weniger als 95% wird als fachlich relevant angesehen. Die Studie liefert das $(1-\alpha)100\%$-Vertrauensintervall $(96\%, 98\%)$ für den Mittelwert der relativen geistigen Leistungsfähigkeit. Welche der folgenden Aussagen trifft zu?

1. Die Abweichung ist statistisch signifikant zum Niveau $\alpha$ und fachlich relevant.
2. Die Abweichung ist statistisch signifikant zum Niveau $\alpha$, aber nicht fachlich relevant.
3. Die Abweichung ist zum Niveau $\alpha$ nicht statistisch signifikant, aber eventuell fachlich relevant.

**2. (7 Punkte)** Maja sucht einen Nachmieter für ihre Wohnung. Es gibt zwei Typen von Leuten, welche sich für Majas Wohnung interessieren: gute und schlechte Bewerber. Während ein guter Bewerber eine Chance von 60% hat, vom Vermieter akzeptiert zu werden, liegt diese beim schlechten Bewerber gerade einmal bei 10%. Insgesamt gehören 70% der Bewerber zur schlechten Sorte und es kann davon ausgegangen werden, dass die Bewerber die Wohnung in zufälliger Reihenfolge besichtigen.

**a)** Leider kann Maja gute Bewerber nicht von schlechten unterscheiden. Wie gross ist die Wahrscheinlichkeit, dass Maja einen schlechten Bewerber vorschlägt und dieser akzeptiert wird?

**b)** Angenommen der Vermieter akzeptiert einen Bewerber. Wie gross ist die Wahrscheinlichkeit, dass es sich dabei um einen schlechten Bewerber handelt?

**c)** Was ist die kleinstmögliche Zahl an Bewerbern, die Maja dem Vermieter vorschlagen muss, damit dieser mit mindestens 95% Wahrscheinlichkeit wenigstens einen davon akzeptiert?

**d)** Maja möchte bei der Nachmietersuche ihr altes, verschlissenes Sofa loswerden. Sie empfiehlt deshalb nur Bewerber, welche bereit sind, das Sofa zu einem Preis von $700 * p$ CHF zu übernehmen, wobei $0 \leq p \leq 1$. Ein schlechter Bewerber kauft das Sofa mit einer Wahrscheinlichkeit von $1 - p/7$, während ein guter nur mit Wahrscheinlichkeit $1 - p$. Wie gross darf $p$ maximal sein, sodass ein Bewerber, der bereit ist das Sofa zu kaufen, mit einer Wahrscheinlichkeit von mindestens $1/8$ ein guter Bewerber ist.

**3. (8 Punkte)** Michael begann heute seinen neuen Job als Kontrolleur beim ZVV. Die Zentrale hat angeordnet, dass die Gruppe von sechs Kontrolleuren, welcher Michael angehört, in der Hauptverkehrszeit zwischen 7:30 und 8:30 fünf Kontrollgänge durchführt. Die Vorschrift der Zentrale für einen solchen Kontrollgang ist, dass alle 6 Kollegen gleichzeitig mit der Kontrolle anfangen, je 10 Fahrgäste kontrollieren, und danach die Gesamtzahl der durch die sechs Kontrolleure gefunden Schwarzfahrer ins Protokoll eintragen. Wir nehmen an, dass alle Kontrollen unabhängig voneinander sind.

**a)** Sei $p$ die Wahrscheinlichkeit, dass ein zufällig ausgewählter Fahrgast Schwarzfahrer ist und $X$ die Gesamtzahl der Schwarzfahrer, die insgesamt von den sechs Kontrolleuren bei je einem Kontrollgang entdeckt werden. Benennen Sie die Verteilung von $X$ und geben Sie die Wahrscheinlichkeit $\mathbb{P}(X = x)$ an für ein beliebiges $x \in \{0, \ldots, 60\}$, und bestimmen Sie $\mathbb{E}(X)$, jeweils in Abhängigkeit von $p$.

**b)** In den vorgeschriebenen fünf Kontrollgängen wurde jeweils folgende Anzahl $X$ von Schwarzfahrern protokolliert

$$x_1 = 2,\ x_2 = 2,\ x_3 = 4,\ x_4 = 1,\ x_5 = 5.$$

Schätzen Sie $p$ mit der Momentenmethode aus der gegebenen Stichprobe.

**c)** Beim ersten Kontrollgang konnte Michael keinen Schwarzfahrer entlarven. Er bemerkt, dass es drei seiner Kollegen genauso ging. Sei $Y$ die Anzahl der Kontrolleure, die beim Kontrollgang keinen Schwarzfahrer finden. Bestimmen Sie die Verteilung von $Y$ in Abhängigkeit von $p$ und berechnen Sie den Maximum-Likelihood Schätzer für $p$ anhand des ersten Kontrollgangs.

**d)** Michael erwischt den ersten Schwarzfahrer im vierten Kontrollgang. Er fragt seine Kollegen, nach der Anzahl $K$ der Fahrgäste die sie kontrollieren mussten bis zum ersten Schwarzfahrer. Die Angaben sind wie folgt:

$$k_1 = 9,\ k_2 = 24,\ k_3 = 10,\ k_4 = 28,\ k_5 = 16,\ k_6 = 31.$$

Dabei bedeutet $k_i = j$, dass bei Kontrolleur $i$ der $j$-te kontrollierte Fahrgast der erste Schwarzfahrer war. Benennen Sie die Verteilung von $K$ und berechnen Sie den Maximum-Likelihood Schätzer für $p$ anhand der Angaben von Michael und seiner Kollegen.

**Bitte wenden!**

**4. (8 Punkte)** Bäckermeister Charles Tüngli besitzt eine Maschine zum Kneten und Partitionieren von Teig, wobei man das Gewicht der gewünschten Partitionen folgendermassen regulieren kann: Laut Herstellerangaben erzeugt die Maschine bei einer Einstellung auf den Wert $\mu$, beliebig viele Teigbällchen, deren Gewichte unabhängig und identisch normalverteilt sind mit Mittelwert $\mu$ und Standardabweichung $\frac{\mu}{100}$. Für seine Spezial-Gipfeli benötigt Charles Teigbällchen zu je 100 Gramm, aber die mittlerweile in die Jahre gekommene Maschine ist schon lange nicht mehr richtig justiert: die produzierten Teigbällchen bei einer Einstellung auf den Wert $\mu$, sind zwar immer noch normalverteilt mit einer Standardabweichung von einem Prozent des eingestellten Wertes, allerdings ist der Mittelwert der Portionen viel zu klein. Anstatt die Maschine neu justieren zu lassen, stellt Charles sie kurzerhand auf einen Wert von 120 Gramm, und möchte nun feststellen, ob die Maschine mit dieser Einstellung denselben Mittelwert liefert wie eine korrekt justierte und auf 100 Gramm eingestellte Maschine.

Ohne jegliche Statistikkenntnisse beschliesst Charles intuitiv folgendes Entscheidungsverfahren: Er misst die Gewichte von 10 zufällig ausgewählten Teigbällchen. Ist deren Mittelwert grösser als 99.5 Gramm und kleiner als 100.5 Gramm, dann behält er diese Einstellung bei, falls nicht, probiert er einen anderen Wert.

**a)** Welcher standardisierte Test aus der Vorlesung ist bei Kenntnis der obigen Angaben am besten geeignet? Begründen Sie Ihre Antwort. Testen Sie dabei ein- oder zweiseitig?

**b)** Auf welchen Parameter testet man? Geben Sie Null- und Alternativhypothese an.

**c)** Geben Sie die Teststatistik von Ihrem Test und die Teststatistik von Charles' Test an. Berechnen Sie den Verwerfungsbereich von Ihrem Test so, dass dieser äquivalent ist zu dem aus Charles' Test.

**d)** Wie gross ist das Niveau $\alpha$ von Charles' Test?