Stochastik
(D-MAVT / D-MATL / RW)

Schreiben Sie stets alle Zwischenschritte und -rechnungen sowie Begründungen auf.

Die für die Aufgaben benötigten Tabellen (Normalverteilung, Quantile der t-Verteilung) wurden mit der Prüfung ausgeteilt.

1. (12 Punkte) Im Hamburger Hafen werden Container auf ein Schiff verladen. Das Gewicht eines Containers in Tonnen ist durch $G = 14X$ gegeben, wobei X eine Zufallsvariable mit Dichte

$$f(x) = \begin{cases} c \cdot x^{10}(1-x)^2, & 0 \le x \le 1; \\ 0, & \text{sonst;} \end{cases}$$

ist.

a) Bestimmen Sie die Konstante c. Geben Sie hierbei Ihren Rechenweg an.

b) Berechnen Sie den Erwartungswert und die Varianz des Gewichtes eines Containers.

Das Schiff wird mit 1500 Containern beladen, wobei angenommen wird, dass die Gewichte der einzelnen Container unabhängig sind und alle dieselbe Verteilung wie die Zufallsvariable G haben.

c) Geben Sie das erwartete Ladegewicht sowie dessen Varianz an.

d) Falls das Gewicht der Ladung 16610 Tonnen übersteigt, so ist das Schiff überladen. Berechnen Sie mittels einer Normalapproximation die Wahrscheinlichkeit dieses Ereignisses.

2. (12 Punkte) In einer Stadt gibt es 20% Holzhäuser. Der Anteil der Betonhäuser ist p. Alle anderen Häuser wurden aus Ziegel gebaut. Aus langjährigen Statistiken weiss man, dass Holzhäuser bei einem Erdbeben der Stärke 6.5 mit Wahrscheinlichkeit 0.3 einstürzen. Betonhäuser stürzen mit Wahrscheinlichkeit 0.9 nicht ein und Ziegelhäuser stürzen mit Wahrscheinlichkeit 0.8 nicht ein.

a) Die Behörde möchte, dass die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Haus bei einem Erdbeben der Stärke 6.5 einstürzt kleiner als 15% ist. Geben Sie das Intervall aller möglichen Werte $p \in [0, 1]$ an, für welche die Forderung der Behörde erfüllt wird.

b) Sei $p = 0.75$. Nach einem Erdbeben der Stärke 6.5 wird der Rettungsdienst zu einem Haus in der Stadt geschickt, wobei man nicht weiss, ob das Haus eingestürzt ist. Mit welcher Wahrscheinlichkeit handelt es sich bei dem Haus um ein eingestürztes Betonhaus?

c) Nach dem Erdbeben stehen Sie vor einem nicht eingestürzten Haus. Mit welcher Wahrscheinlichkeit handelt es sich um ein Betonhaus, falls p wiederum 0.75 ist?

d) N_1 und N_2 seien die Anzahl Erdbeben der Jahre 2007 resp. 2008. Wir nehmen an, dass N_1 und N_2 unabhängig und Poisson-verteilt sind mit Parameter $\lambda = 0.1$. Wie gross ist die Wahrscheinlichkeit, dass in den Jahren 2007-2008 insgesamt 2 oder mehr Erdbeben auftreten?

Bitte wenden!

3. (12 Punkte) Der Patissier Giorgio hat ein altes Rezept für Butterkekse im Schrank seiner Grossmutter gefunden und hat es sofort ausprobiert. Für die neuen Kekse hat er so viele Komplimente von seinen Kunden bekommen, dass er sich entschlossen hat, diese im grossen Stil zu verkaufen. Dazu braucht er aber eine neue Maschine. Ihm stehen zwei Maschinen A und B zur Auswahl. Nach langem Abwägen von Vor- und Nachteilen der einzelnen Maschinen, kommt er zum Schluss, dass er die Maschine kauft, welche in einer Stunde am meisten produzieren kann. Man lässt die beiden Maschinen 10 mal jeweils eine Stunde Kekse produzieren und man erhält folgende Werte:

Stunde Nr.	1	2	3	4	5	6	7	8	9	10
Maschine A: x_i	995	970	955	1005	980	975	980	985	1000	965
Maschine B: y_i	985	1020	975	1005	1000	990	1000	1010	985	1015

$\bar{x}_{10} = 981,\ s_x^2 = 248.89,\ \bar{y}_{10} = 998.5,\ s_y^2 = 211.39,\ s_{x-y}^2 = 479.17,\ s_{pool}^2 = 230.14.$

Man kann davon ausgehen, dass die Anzahl Kekse pro Stunde gut durch eine Normalverteilung approximiert wird. Führen Sie einen geeigneten t-Test durch, um zu sehen, ob sich die Produktionsmengen pro Stunde der beiden Maschinen signifikant unterscheiden.

a) Wie lautet die Nullhypothese?

(i) Maschine A produziert mehr Kekse in einer Stunde als Maschine B.
(ii) Maschine A produziert gleichviele Kekse in einer Stunde wie Maschine B.
(iii) Maschine A produziert weniger Kekse in einer Stunde als Maschine B.
(iv) Die Maschinen A und B produzieren unterschiedlich viele Kekse pro Stunde.

b) Wie lautet die korrekte Alternativ-Hypothese?

(i) Maschine A produziert mehr Kekse in einer Stunde als Maschine B.
(ii) Maschine A produziert weder mehr noch weniger Kekse in einer Stunde als Maschine B.
(iii) Maschine B produziert mehr Kekse in einer Stunde als Maschine A.
(iv) Entweder Maschine A oder B produziert im Schnitt mehr Kekse pro Stunde als die andere Maschine.

c) Wie muss der zugehörige Test durchgeführt werden?

(i) Einseitig und gepaart.
(ii) Einseitig und ungepaart.
(iii) Zweiseitig und gepaart.
(iv) Zweiseitig und ungepaart.

Führen Sie nun gemäss Ihren Antworten bei den Aufgaben a), b) und c) den geeigneten t-Test zum Niveau 5% durch.

d) Geben Sie den Verwerfungsbereich an! Kann Giorgio nach diesem Test eine eindeutige Kaufentscheidung fällen? Falls ja, wie entscheidet er sich? Begründen Sie die Antworten!

4. (12 Punkte) Die Wahrscheinlichkeit, dass an einem beliebigen Tag in einer Grossstadt erhöhte (d.h. einen kritischen Wert überschreitende) Schadstoffemissionen gemessen werden sei p. Zusätzlich nehmen wir an, dass die Emissionshöhen an unterschiedlichen Tagen unabhängig sind.

a) Sei in einem Zeitraum von n Tagen X_n die Anzahl derjenigen Tage, an denen erhöhte Emissionswerte gemessen werden. Welche Verteilung hat X_n?

b) Innerhalb der letzten 365 Tage wurden an 219 Tagen erhöhte Emissionen gemessen. Finden Sie eine vernünftige Schätzung für p und bestimmen Sie unter Verwendung der Normalapproximation ein 99%-Vertrauensintervall für p.

Neue verkehrspolitische Massnahmen werden implementiert, von denen man sich eine Reduktion von p auf unter 0.4 erhofft. 100 Tage nach Implementierung dieser Massnahmen soll daher versucht werden, statistisch zu belegen, dass p auf unter 0.4 gefallen ist.

c) Formulieren Sie die für die Belegung letzterer Aussage geeigneten Null- und Alternativhypothesen und bestimmen Sie mittels Normalapproximation den Verwerfungsbereich für das Signifikanzniveau 0.05.

d) Während dieses Zeitraums von 100 Tagen nach Implementierung der Massnahmen werden an 28 Tagen erhöhte Emissionen gemessen. Erklären Sie kurz das Resultat des Tests aus Aufgabenteil c) und bestimmen Sie mittels Normalapproximation den relevanten P-Wert.