Dr. Rudolf Riedi
HTA-FR, 2017-18

Statistik
Vorbereitung Test 1

Buch, Notizen und Formelsammlung erlaubt.

Der Taschenrechner kann zur Überprüfung benützt werden, und zur Berechnung von numerischen Ausdrücken. Die Herleitung muss ersichtlich sein! (Ein Resultat ohne Berechnungsschritte wird mit Null Punkten bewertet).

In den Multiple-Choice-Fragen kreuze man *jedes* Feld an, die als richtig befundenen Felder mit einem Kreis, die als falsch befundenen Felder mit einem Kreuz.

Name ....

---

1. (3 Punkte)
   Eine Stichprobe enthält die folgenden Daten:

   $$3 \quad 4 \quad 2 \quad 1 \quad 6 \quad 1 \quad 4 \quad 5 \quad 4$$

   Berechnen Sie: Modus, Median, arithmetisches Mittel, Varianz, 1. und 3. Quartil.

   Lösung

   Die Daten sortiert:

   $$1 \quad 1 \quad 2 \quad 3 \quad 4 \quad 4 \quad 4 \quad 5 \quad 6$$

   Wir finden

   - Modus: 4
   - Median: 4
   - Mittel: 30/9=3.33
   - Varianz: 3
   - Standard Abweichung: 1.732
   - 1.Quartil: 2
   - 3.Quartil: 4

   Für die Quartile beachte man, dass die Daten nicht klassiert sind. Man wählt also einen Datenpunkt als Quartil aus.

2. (3 Punkte) Im folgenden bezeichnet A ein Ereignis mit $0 < P[A] < 1$. Entscheiden Sie, ob die folgenden Aussagen richtig oder falsch sind.

   $\boxed{A \text{ und } \bar{A} \text{ sind unabhängig}}$ $\boxed{A \text{ und } \bar{A} \text{ sind abhängig}}$ $\boxed{A \text{ und } \bar{A} \text{ können abhängig oder unabhängig sein}}$

   Richtig: $\boxed{A \text{ und } \bar{A} \text{ sind abhängig}}$

   Begründung 1: Wenn man weiss, dass A eintrifft, dann weiss man auch dass A^c nicht eintrifft. Also sind sie abhängig.

   Begründung 2: $P[A \cap A^c] = P[\emptyset] = 0$. Aber $P[A] \cdot P[A^c] \neq 0$ denn sonst wäre entweder $P[A] = 0$ oder $P[A] = 1$. Also ist $P[A \cap A^c] \neq P[A] \cdot P[A^c]$.

3. (6 Punkte) Eine Erhebung der Wartezeiten an einem Server unterscheidet zwei Ankunkftszeiten (früh = zwischen 8 und 10 Uhr, spät = zwischen 10 und 12 Uhr) sowie zwei Wartezeit (kurz = weniger als 3 Sekunden, lang = mehr als 3 Sekunden).

   Es zeigt sich folgendes Ergebnis: 350 Kunden kamen früh, 800 im Ganzen. 550 Kunden erfuhren kurze Wartezeiten, 250 davon kamen früh.

   Lösung

   Man stellt zuerst die Kontigenztafel auf. Diese lautet

	kurz	lang	total
früh	250	100	350
spät	300	150	450
total	550	250	800

   (a) Wieviele Kunden kamen spät und erfuhren eine lange Wartezeit? 150

   (b) Falls man einen der 800 Kunden zufällig wählen würde, wie gross ist die Wahrscheinlichkeit, dass er eine lange Wartezeit erfuhr? 250/800=5/16

   (c) Für einen zufällig gewählten späten Kunden, wie gross ist die Wahrscheinlichkeit, dass er eine lange Wartezeit erfuhr? 150/450=1/3

4. (a) (2 Punkte) Vervollständigen Sie (A und B sind beliebig) : $P[A \cap B] = P[A] * P[\ldots\ldots\ldots\ldots\ldots]$. Antwort: $P[A \cap B] = P[A] * P[B|A]$.

   (b) (2 Punkte) Betrachten Sie die Formel $P[A|B] = 1 - P[A|\overline{B}]$ und entscheiden Sie:

   $\boxed{\text{Formel ist richtig}}$ $\boxed{\text{Formel ist falsch}}$

   Antwort: $\boxed{\text{Formel ist falsch}}$

   Begründung: Nehmen wir als einfaches Beispiel den Fall, dass A und B unabhängig sind. Dann sind auch A und B^c unabhängig und dann müsste also gelten: $P[A] = 1 - P[A]$, also $P[A] = 1/2$. Das ist natürlcih nicht immer so. Also ist die Formel im Allgemeinen falsch. Sie kann in speziellen Situationen zufällig stimmen.

5. (6 Punkte)

   In einer Stadt sind 65% der Bewohner Anhänger von Mr. Obama. In einem Zugabteil befinden sich drei Bewohner dieser Stadt, die völlig zufällig zusammengekommen sind.

   (a) (3 Punkte) Wie gross ist die Wahrscheinlichkeit, dass die Anhänger von Mr. Obama unter diesen Dreien in der Überzahl sind, wenn die Stadt 40'000 Einwohner hat? Man rechne exakt.

   (b) (3 Punkte) Berechnen Sie eine Annäherung dieser Wahrscheinlichkeit, falls die Einwohnerzahl nicht bekannt, aber sehr gross ist.

   Lösung

   (a) Dies ist eine Stichprobe (ohne Zurücklegen) vom Umfang 3, aus 40'000=, wobei 40'000*0.65=26'000 das Merkmal tragen. Die Anzahl X von Obama-Anhängern unter den Dreien folgt also der Hypergeometrischen Verteilung $X \sim H(40'000; 26'000; 3)$.

   Wir sind interessiert an der Wahrscheinlichkeit $P[X = 2] + P[X = 3] = 1 - P[X = 0] - P[X = 1]$. Die zweite variante ist schneller zu rechnen. Wir zeigen beide:

$$P[X = 2] = \frac{\binom{26'000}{2} \cdot \binom{14'000}{1}}{\binom{40'000}{3}} = \frac{26000 * 25999 * 14000}{40000 * 39999 * 39998} \cdot \frac{3!}{1!2!} = 0.4436$$

$$P[X=3]=\frac{\binom{26'000}{3}\cdot\binom{14'000}{0}}{\binom{40'000}{3}}=\frac{26000*25999*25998}{40000*39999*39998}\cdot\frac{3!}{3!0!}=\frac{26000}{40000}\frac{25999}{39999}\frac{25998}{39998}=0.2746$$

Die gesamte Wahrscheinlichkeit beträgt also $0.4436+0.2746=0.7183$

Zweiter Weg

$$P[X=0]=\frac{\binom{26'000}{0}\cdot\binom{14'000}{3}}{\binom{40'000}{3}}=\frac{14000*13999*13998}{40000*39999*39998}\cdot\frac{3!}{0!3!}=\frac{14000}{40000}\frac{13999}{39999}\frac{13998}{39998}=0.0429$$

$$P[X=1]=\frac{\binom{26'000}{1}\cdot\binom{14'000}{2}}{\binom{40'000}{3}}=\frac{26000*14000*13999}{40000*39999*39998}\cdot\frac{3!}{1!2!}=0.2389$$

Die gesamte Wahrscheinlichkeit beträgt also $1-0.0429-0.2389=0.7183$ wie oben.

(b) Falls die Anzahl Einwohner unbekannt ist, betrachten wir die Stichprobe näherungsweise als Bernoulli-Experiment, genauer gesagt, wir ziehen zufällig einen Einwohner und legen ihn zurück (bie einer Stichprobe legt man den Einwohner nicht zurück). Dieses Ziehen wird drei mal wiederholt.

Wir haben also einen Baum mit drei identischen Versuchen, jeweils mit Erfolgswahrscheinlichkeit $p=65\%=0.65$.

Die Wahrscheinlichkeit, dass genau 3 Ziehungen erfolgreich waren, entspricht dem Ast EEE (E=Erfolg=Anhänger von Obama) und beträgt $p\cdot p\cdot p=p^3$. Alternative: Als Binomial-Verteilung gerechnet möchten wir genau 3 erfolge unter den 3 Versuchen; die W'keit beträgt

$$P[3\text{ Anhänger}]=\binom{3}{3}p^3=0.65^3=0.2746$$

Die Wahrscheinlichkeit, dass genau 2 Ziehungen erfolgreich waren, also genau 2 unter den 3 Einwohnern Anhänger von Obama sind, entspricht den Ästen EEN, ENE und NEE (N=Nicht-Erfolg), von denen jeder die W'keit $p\cdot p\cdot(1-p)$ besitzt. Wir halten also, wie bei der Binomial-Verteilung

$$P[2\text{ Anhänger}]=\binom{3}{2}p^2(1-p)^1=3\cdot 0.65^2\cdot 0.35=0.4436$$

Die gesuchte Wahrscheinlichkeit ist die Summe der beiden obigen, also $0.7182=71.82\%$.

6. (3 Punkte) Ein Bewerber wird auf seine Eignung getestet. Unter den Fragen befindet sich eine Multiple-Choice Frage mit 4 möglichen Antworten, von denen aber nur genau eine richtig ist.

Ein fähiger Bewerber wird die Frage richtig beantworten. Ein unfähiger Bewerber wird eine der 4 möglichen Antworten zufällig auswählen. Aus Erfahrung weiss man, dass 80% der Bewerber fähig sind.

Wie gross ist die Wahrscheinlichkeit, dass ein Bewerber, der richtig auf die Multiple-Choice Frage geantwortet hat, wirklich fähig ist.

Hinweis: Bayes' Theorem. A='Bewerber unfähig', B='Frage richtig beantwortet.'

Lösung

(a) Gesucht ist $P[A^c|B]=1-P[A|B]$.

Aus der Aufgabenstellung wissen wir, dass $P[A]=20\%$ und $P[B|A]=1/4$. Mittels Bayes:

$$P[A|B]=\frac{P[B|A]P[A]}{P[B|A]P[A]+P[B|A^c]P[A^c]}=\frac{0.25\cdot 0.20}{0.25\cdot 0.20+1\cdot 0.80}=\frac{0.05}{0.85}=\frac{1}{17}=0.0588$$

Damit wird die gesuchte Wahrscheinlichkeit: **0.9412**