Abitur Mathematik: Musterlösung

Aufgabe 8: Stochastik (WTR)

Nordrhein-Westfalen 2013

Aufgabe 8

a)

(1) WAHRSCHEINLICHKEIT FÜR „KEINE ANGABE“ ERMITTELN

Nach der Laplace-Formel ist

$$P(\text{keine Angabe}) = \frac{\text{Anzahl der Personen, die keine Angabe machten}}{\text{Gesamtzahl der Befragten}}$$

Die Zahl der befragten Personen, die keine Angabe gemacht haben ist die Summe der in der Zeile „keine Angabe“ aufgelisteten Ergebnisse, also $12 + 14 + 8 + 4 + 24 = 62$. Insgesamt wurden 1005 Personen befragt. Somit ist

$$P(\text{keine Angabe}) = \frac{62}{1005} \approx 6{,}17\ \%.$$

ERGEBNIS

Die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person aus den 1005 Umfrageteilnehmern keine Angabe gemacht hat, beträgt etwa 6,2 %.

(2) ANTEIL NICHT-SCHÜLER MIT ANGABE „JA“ ERMITTELN

Die Aufgabenstellung ist zweideutig, je nachdem, ob der Nebensatz die Grundgesamtheit einschränkt, oder sich nur auf den zu ermittelnden Anteil bezieht.

1. Interpretation:

„Bestimmen Sie unter den 14- bis 29-Jährigen, die mit „Ja“ geantwortet haben, den Anteil der Nicht-Schüler.“

2. Interpretation:

„Bestimmen Sie unter allen 14- bis 29-Jährigen den Anteil der Nicht-Schüler, die mit „Ja“ geantwortet haben.“

Zwei Drittel der 57 Schüler haben mit „Ja" gestimmt, das sind 38 Stück. Da insgesamt 166 Personen der Altersklasse 14-29 Jahre mit „Ja" gestimmt haben, waren darunter $166 - 38 = 128$ Nicht-Schüler.

Die Grundgesamtheit ist bei der 1. Interpretation die Menge aller 14- bis 29-jährigen, die mit „Ja" gestimmt haben, also laut Tabelle 166 Personen. Hier ist also der zu berechnende Anteil $\frac{128}{166} = \frac{64}{83} \approx 77{,}1\ \%$.

Bei der 2. Interpretation ist die Grundgesamtheit die Menge aller befragten 14- bis 29-jährigen, also laut Tabelle 211 Personen. In diesem Fall ist der zu berechnende Anteil $\frac{128}{211} \approx 60{,}7\ \%$.

b)

(1) BEDINGTE WAHRSCHEINLICHKEIT FÜR HERKUNFT WESTDEUTSCHLAND ERMITTELN

Bezeichnungen:

J = der Befragte stimmt mit „Ja"
$\bar{J}$ = der Befragte stimmt *nicht* mit „Ja"
W = der Befragte stammt aus Westdeutschland
O = der Befragte stammt aus Ostdeutschland

Gesucht ist die Wahrscheinlichkeit des Ereignisses W unter der Bedingung J, also

$$P_J(W) = \frac{P(W \cap J)}{P(J)}$$

Dabei ist

$$P(W \cap J) = \frac{\text{Anzahl der Westdeutschen, die mit "Ja" gestimmt haben}}{\text{Gesamtzahl der Teilnehmer}}$$
$$= \frac{643}{1005} \text{ und}$$

$$P(J) = \frac{\text{Anzahl der Befragten, die mit "Ja" gestimmt haben}}{\text{Gesamtzahl der Teilnehmer}}$$
$$= \frac{792}{1005}\text{, also}$$

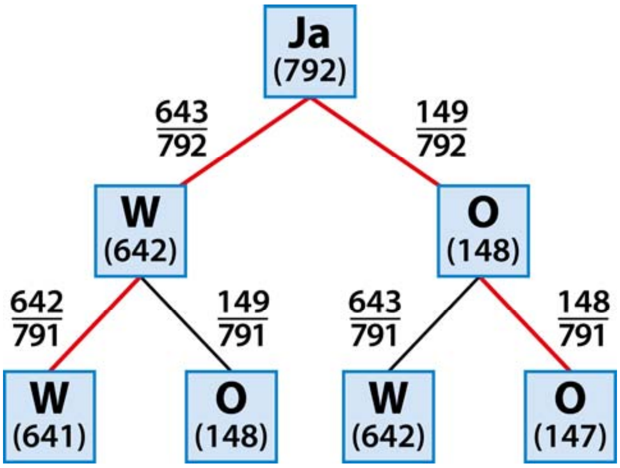
$$P_J(W) = \frac{\frac{643}{1005}}{\frac{792}{1005}} = \frac{643}{792} \approx 81{,}2\ \%.$$

(2) WAHRSCHEINLICHKEIT FÜR ZWEI TEILNEHMER AUS DEMSELBEN LANDESTEIL ERMITTELN

Aus den 792 Personen, die mit „Ja" gestimmt haben, wird zweimal ohne Zurücklegen gezogen. Bei der ersten Ziehung wird mit Wahrscheinlichkeit $\frac{643}{792}$ ein Westdeutscher und mit Wahrscheinlichkeit $\frac{149}{792}$ ein Ostdeutscher gewählt. Bei der zweiten Ziehung stehen nur noch 791

Personen zur Auswahl, davon entweder 643 oder 642 aus Westdeutschland, je nach Herkunft des ersten Befragten.

Baumdiagramm:



Die rot gekennzeichneten Pfade sind diejenigen, die zum Ereignis „Die zweite Person stammt aus demselben Teil Deutschlands wie die erste" beitragen. Die Wahrscheinlichkeit dieses Ereignisses ist nach der 2. Pfadregel

$$\begin{aligned} P(WW) + P(OO) &= P_W(W) + P_O(O) \\ &= \frac{643}{792} \cdot \frac{642}{791} + \frac{149}{792} \cdot \frac{148}{791} \\ &\approx 0{,}659 + 0{,}035 = 0{,}694. \end{aligned}$$

ERGEBNIS

Die Wahrscheinlichkeit, dass beide zufällig ausgewählten Personen aus derselben Region stammen, wenn beide mit „Ja" gestimmt haben, beträgt 69,4 %.

c)

(1) BINOMIALVERTEILUNG BEGRÜNDEN

Voraussetzungen für eine Binomialverteilung:

- Die n befragten Personen antworten unabhängig voneinander.
- Es werden jeweils nur zwei Antworten unterschieden: „Nein" oder „nicht Nein" (d. h. die Antwortmöglichkeiten „Ja" und „keine Angabe" werden zusammengefasst).
- Die Wahrscheinlichkeit, dass ein Teilnehmer mit „Nein" antwortet, bleibt bei allen n Befragungen gleich.

Unter diesen drei Bedingungen handelt es sich um n unabhängige Durchführungen desselben Bernoulli-Experiments, also ist die Anzahl der Treffer binomialverteilt.

Im vorliegenden Fall wird die Antwort „Nein" als Treffer angesehen. Die ersten zwei Bedingungen können ohne weiteres als erfüllt angenommen werden.

Die dritte Bedingung ist streng genommen nicht erfüllt, da dieselbe Person nicht mehrfach befragt wird, d. h. das passende Modell wäre „Ziehen ohne Zurücklegen" (mit Beachtung der Reihenfolge), während die Binomialverteilung von „Ziehen mit Zurücklegen" ausgeht. Sofern aber n sehr viel kleiner als die Anzahl der befragten 50- bis 59-jährigen ist, ändert sich die Trefferwahrscheinlichkeit p nur geringfügig:

$$p = \frac{\#\mathrm{N}}{\#\mathrm{G}},$$

wobei #N die Anzahl der noch zur Auswahl stehenden (also noch nicht befragten) 50- bis 59-jährigen ist, die mit „Nein" stimmen, und #G die Gesamtzahl der noch nicht befragten 50- bis 59-jährigen. Es ist

$0 \leq \#\mathrm{N} \leq 19$ und $152 - n \leq \#\mathrm{G} \leq 152$.

Somit gilt z. B. für $n \leq 19$

$$0 \leq \frac{19-n}{152} \leq p \leq \frac{19}{152-n} \leq \frac{1}{7}.$$

Anfangs ist laut Tabelle $\#\mathrm{N} = 19$ und $\#\mathrm{G} = 152$, also

$$p = \frac{\#\mathrm{N}}{\#\mathrm{G}} = \frac{19}{152} = \frac{1}{8}.$$

Für kleine n ist also die Binomialverteilung eine gute Näherung für die Verteilung von X.

(2) WAHRSCHEINLICHKEIT FÜR $X = \mu$ BERECHNEN

Gemäß Teilaufgabe (1) nehmen wir näherungsweise an, dass X binomialverteilt ist mit Parametern $n = 40$ und $p = \frac{1}{8}$. Demzufolge ist der Erwartungswert $\mu = n \cdot p = 40 \cdot \frac{1}{8} = 5$.

Gesucht ist also

$$P(X = 5) = \binom{40}{5} \cdot \left(\frac{1}{8}\right)^5 \cdot \left(\frac{7}{8}\right)^{35} \approx 0{,}188$$

laut Bernoulliformel.

ERGEBNIS

Die Wahrscheinlichkeit dafür, dass unter 40 zufällig ausgewählten Personen $\mu = 5$ Personen mit „Nein" antworten, ist etwa 18,8 %.

Bemerkung:

In Wirklichkeit ist X hypergeometrisch verteilt zu den Parametern $n = 40$, $m = 19$ und $N = 152$, wobei auch hier der Erwartungswert $\mu = 5$ ist. Es ist $P(X = 5) = \frac{\binom{n}{5} \cdot \binom{N-n}{m-5}}{\binom{N}{m}} = \frac{\binom{40}{5} \cdot \binom{112}{14}}{\binom{152}{19}} \approx 0{,}218$.

(3) WAHRSCHEINLICHKEIT FÜR $X \leq 8$ BERECHNEN

Es stimmen genau dann mindestens 42 der 50 Befragten nicht mit „Nein", wenn höchsten 8 von ihnen mit „Nein" stimmen. Gefragt ist also $P(X \leq 8)$, was in Tabelle 3 (kumulierte Binomialverteilung für $n = 50$) in der Zeile $k = 8$ und der Spalte $p = 0{,}125$ nachgeschlagen werden kann: Es ist $P(X \leq 8) \approx 0{,}8339$.

ERGEBNIS

Die Wahrscheinlichkeit dafür, dass unter 50 zufällig ausgewählten Teilnehmern mindestens 42 mit „Nein" stimmen, beträgt etwa 83,4 %.

Bemerkung:

In Wirklichkeit ist $P(X \leq 8) \approx 0{,}879$.

(4) WAHRSCHEINLICHKEIT FÜR 10 ANGABEN BERECHNEN

Es wurden $211 + 150 + 191 = 552$ Personen zwischen 14 und 49 Jahren befragt. Davon haben $12 + 14 + 8 = 34$ keine Angabe gemacht, d. h. $552 - 34 = 518$ Personen haben mit „Ja" oder „Nein" gestimmt.

Der Pfad, dessen Wahrscheinlichkeit zu bestimmen ist, zeichnet sich dadurch aus, dass bei jeder Befragung mit „Ja" oder „Nein" gestimmt wird.

Bei der ersten Befragung wird also mit Wahrscheinlichkeit $p_1 = \frac{518}{552}$ eine Angabe gemacht. Mit jeder Befragung wird die Zahl der noch zur Auswahl stehenden Teilnehmer, die eine Angabe machen (Zähler der Wahrscheinlichkeit), um eins kleiner, ebenso die Gesamtzahl der noch zur Auswahl stehenden Teilnehmer (Nenner der Wahrscheinlichkeit).

Nach der 1. Pfadregel ist also

$$P(10\text{-mal Angabe}) = \frac{518}{552} \cdot \frac{517}{551} \cdot \frac{516}{550} \cdot \frac{515}{549} \cdot \frac{514}{548} \cdot \frac{513}{547} \cdot \frac{512}{546} \cdot \frac{511}{545} \cdot \frac{510}{544} \cdot \frac{509}{543} \approx 0{,}527.$$

ERGEBNIS

Mit einer Wahrscheinlichkeit von etwa 52,7 % stimmen die ersten 10 Befragten mit „Ja" oder „Nein".

d)

(1) BEDEUTUNG DES KONFIDENZINTERVALLS ERLÄUTERN

Angenommen, es werden n Männer für die Stichprobe zufällig ausgewählt (wobei auch mehrfache Befragung derselben Person zugelassen wird), von denen k^* das Verbot von Gen-Mais befürworten. Dann ist k^* ein möglicher Wert einer binomialverteilten Zufallsvariablen mit Parametern n und p_M.

Der nächstliegende Schätzwert für p_M angesichts der Stichprobe ist die Anteil p_M^* der Befürworter in der Stichprobe, aber trotzdem ist es sehr unwahrscheinlich, dass $p_M = p_M^*$ ist. Deswegen versucht man anhand der Stichprobe statt eines einzigen Schätzwertes einen Bereich anzugeben, in dem sich p_M sehr wahrscheinlich befindet.

Ist dieser Bereich ein Intervall, das nach einer festen Regel nur in Abhängigkeit vom zufälligen Wert k^* aus der Stichprobe ermittelt wird, so handelt es sich um ein Konfidenzintervall. Ist die Regel so geartet, dass sie mit einer Wahrscheinlichkeit von mindestens 95 % ein Intervall liefert, das den wahren Wert p_M enthält, so ist das nach dieser Regel zu einem vorgegebenen Stichprobenergebnis k^* bestimmte Intervall ein 95 %-Konfidenzintervall für p_M. In diesem Sinne kann man sagen, dass der zu schätzende Anteil der Gen-Mais-Gegner unter den Männern mit einer Wahrscheinlichkeit von mindestens 95 % in K_M liegt.

(2) 95 %-KONFIDENZINTERVALL FÜR p_F BESTIMMEN

Das 95 %-Konfidenzintervall K_F für p_F besteht aus allen p, für die der 95 %-Annahmebereich der Binomialverteilung zu den Parametern $n = 518$ und p die beobachtete Zahl 426 aus der Stichprobe enthält. Dabei ist der 95 %-Annahmebereich ein um den Erwartungswert $\mu_p = n \cdot p$ zentriertes Intervall $I_p = [\mu_p - m_p; \mu_p + m_p]$ mit der Eigenschaft, dass eine Zufallsvariable $Z \sim B(n; p)$ mit einer Wahrscheinlichkeit von mindestens 95 % Werte in I_p annimmt.

Laut den in Tabelle 1 angegebenen σ-Regeln ist der 95 %-Annahmebereich der Binomialverteilung mit Parametern n und p näherungsweise das Intervall $[n \cdot p - 1{,}96 \cdot \sigma; n \cdot p + 1{,}96 \cdot \sigma]$, wobei $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ die Standardabweichung der Binomialverteilung zu den Parametern n und p ist. Die Untergrenze p_1 und die Obergrenze p_2 von K_F können also näherungsweise über die Gleichung

$$n \cdot p \pm 1{,}96 \cdot \sigma = 426$$

berechnet werden.

Durch Einsetzen von $n = 518$ und $\sigma = \sqrt{518 \cdot p \cdot (1-p)}$ ergibt sich

$$518 \cdot p \pm 1{,}96 \cdot \sqrt{518 \cdot p \cdot (1-p)} = 426$$

$$\pm 1{,}96 \cdot \sqrt{518 \cdot p \cdot (1-p)} = 426 - 518 \cdot p$$

$1{,}96^2 \cdot 518 \cdot p \cdot (1-p) = (426 - 518 \cdot p)^2$

$1989{,}9488 \cdot (p - p^2) = 426^2 - 2 \cdot 426 \cdot 518 \cdot p + 518^2 p^2$

$1989{,}9488 \cdot (p - p^2) = 181476 - 441336p + 268324p^2$

$1989{,}9488 \cdot (p - p^2) = 181476 - 441336p + 268324p^2$

$(268324 + 1989{,}9488)p^2 - (441336 + 1989{,}9488)p + 181476 = 0$

$270313{,}9488p^2 - 443325{,}9488p + 181476 = 0$

Mit Hilfe der quadratischen Lösungsformel erhält man

$$p_1 = \frac{443325{,}9488 - \sqrt{443325{,}9488^2 - 4 \cdot 270313{,}9488 \cdot 181476}}{2 \cdot 270313{,}9488}$$
$$\approx 0{,}7871$$

und

$$p_2 = \frac{443325{,}9488 + \sqrt{443325{,}9488^2 - 4 \cdot 270313{,}9488 \cdot 181476}}{2 \cdot 270313{,}9488}$$
$$\approx 0{,}8529.$$

ERGEBNIS

Somit ist $K_F = [0{,}7871; 0{,}8529]$ näherungsweise ein 95 %-Konfidenzintervall für p_F.

Bemerkung:

Ohne Näherung mittels σ-Formeln erhält man als 95 %-Konfidenzintervall $K_F = [0{,}7867; 0{,}8544]$.