Abitur Mathematik: Musterlösung

Prüfungsteil 2, Aufgabe 8 Stochastik

Nordrhein-Westfalen 2012 LK

Aufgabe a (1) und (2)

1. SCHRITT: VERTEILUNG ANGEBEN

Da die Anzahl der Wahlberechtigten sehr viel größer ist, als die betrachteten Stichproben (höchstens 200), kann näherungsweise davon ausgegangen werden, dass die Zufallsgröße X jeweils binomialverteilt ist mit Parametern n (Anzahl der zufällig ausgewählten Wahlberechtigten) und p (Wahrscheinlichkeit, dass ein zufällig ausgewählter Wahlberechtigte die genannte Partei wählen würde).

(1): $n = 100 \quad p = 0{,}25 \quad k \geq 30$

(2): $n = 200 \quad p = 0{,}1 \quad 8 \leq k \leq 18$

2. SCHRITT: WAHRSCHEINLICHKEITEN ERMITTELN

Bei (1) ist der Stichprobenumfang $n = 100$, bei (2) ist $n = 200$. Mit Hilfe der kumulativen Verteilungstabellen ergibt sich

$P(X \geq 30) = 1 - P(X \leq 29) \approx 1 - 0{,}8505 = 0{,}1495$ für (1)

und

$P(8 \leq X \leq 18) = P(X \leq 18) - P(X \leq 7) \approx 0{,}3724 - 0{,}0005 = 0{,}3719$ für (2).

Aufgabe a (3)

1. SCHRITT: BEDINGUNG FÜR NORMALAPPROXIMATION

Um die Binomialverteilung mit der Normalverteilung sinnvoll approximieren zu können, muss für die Standardabweichung gelten: $\sigma > 3$ (Laplace – Bedingung).

2. SCHRITT: μ UND σ BESTIMMEN

Bei (2) ist X binomialverteilt zu den Parametern $n = 200$ und $p = 0{,}1$, also ist der Erwartungswert $\mu = n \cdot p = 20$ und die Standardabweichung $\sigma = \sqrt{n \cdot p \cdot (1-p)} \approx 4{,}24 > 3$.

3. SCHRITT: WAHRSCHEINLICHKEIT BERECHNEN

$$\begin{aligned} P(8 \leq X \leq 18) &\approx \Phi\left(\frac{18-20+0{,}5}{4{,}24}\right) - \Phi\left(\frac{8-20-0{,}5}{4{,}24}\right) \\ &\approx \Phi(-0{,}35) - \Phi(-2{,}95) \\ &= \Phi(2{,}95) - \Phi(0{,}35) \ (\text{wegen } \Phi(-z) = 1 - \Phi(z)) \\ &\approx 0{,}9984 - 0{,}6368 \\ &= 0{,}3616. \end{aligned}$$

Aufgabe a (4)

SITUATION MODELLIEREN

Annahmen:

- Es werden n Wahlberechtigte unabhängig voneinander befragt.
- Für jede Befragung gibt es genau zwei mögliche Ausgänge: „Befragter würde CDU wählen" oder „Befragter würde nicht CDU wählen".
- Die Wahrscheinlichkeit dafür, dass ein Befragter die CDU wählen würde, beträgt bei jeder Befragung 36 %.

Die Anzahl X der CDU-Wähler unter den Befragten ist somit binomialverteilt zu den Parametern n und $p = 0{,}36$.

Gesucht ist das kleinste $n \in \mathbb{N}$ mit der Eigenschaft, dass

$P(X \geq 1) \geq 90\ \%$ gilt. Dabei ist:

$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{n}{0} \cdot 0{,}36^0 \cdot 0{,}64^n = 1 - 0{,}64^n$.

Die zu erfüllende Ungleichung lautet also

$1 - 0{,}64^n \geq 0{,}9$	$\mid -0{,}9\ + 0{,}64^n$
$0{,}64^n \leq 0{,}1$	$\mid$ logarithmieren
$\ln(0{,}64^n) \leq \ln(0{,}1)$	$\mid$ Logarithmengesetz anwenden
$n \cdot \ln(0{,}64) \leq \ln(0{,}1)$	$\mid : \ln(0{,}64)$
$n \geq \frac{\ln(0{,}1)}{\ln(0{,}64)} \approx 5{,}16$	

Die kleinste natürliche Zahl mit dieser Eigenschaft ist 6.

Um mit 90 %iger Wahrscheinlichkeit mindestens einen CDU-Wähler zu finden, müssen mindestens 6 Wahlberechtigte befragt werden.

Aufgabe a (5)

1. SCHRITT: WAHRSCHEINLICHKEIT BERECHNEN

Die Wahrscheinlichkeit, dass in einer fest vorgegebenen Reihenfolge genau drei von sieben angerufenen Wahlberechtigten CDU-Wähler sind, beträgt

$0{,}36^3 \cdot 0{,}64^4$.

Es gibt fünf Reihenfolgen, bei denen die drei CDU-Wähler direkt hintereinander befragt werden (die Position des 3. CDU-Wählers kann bei 3, 4, 5, 6 oder 7 sein). Daher ist die gesuchte Wahrscheinlichkeit

$5 \cdot 0{,}36^3 \cdot 0{,}64^4 \approx 0{,}039 = 3{,}9\ \%$.

Aufgabe b

1. SCHRITT: 95 %-KONFIDENZINTERVALL FÜR p_G BESTIMMEN

Sei p_G der Anteil der Wähler, die sich am nächsten Sonntag für die Grünen entscheiden würden. Sei ferner X die Anzahl der Grünen-Wähler unter den 1352 Befragten. X wird als binomialverteilt zu den Parametern $n = 1352$ und p_G angenommen.

Dementsprechend ist der Erwartungswert von X

$\mu = \mathrm{E}(X) = n \cdot p_G = 1352 \cdot p_G$

und sofern die Laplace-Bedingung erfüllt ist, gilt laut Tabelle 1 (σ-Regeln)

$P(\mu - 1{,}96\sigma \leq X \leq \mu + 1{,}96\sigma) \approx 0{,}95$, d. h.

$P(|X - \mu| \leq 1{,}96\sigma) \approx 0{,}95 \Leftrightarrow P\left(\left|\frac{X}{n} - p_G\right| \leq \frac{1{,}96\sigma}{n}\right) \approx 0{,}95.$

Gesucht ist ein Teilintervall $K_G \subset [0; 1]$, das (in einem geeigneten Sinne) mit mindestens 95 % Wahrscheinlichkeit p_G enthält. Gemeint ist: K_G enthält genau die Schätzwerte $p \in [0; 1]$, für die das Ereignis $\frac{X}{n} \approx 10\ \%$, das bei der Umfrage eingetreten ist, die Bedingung $\left|\frac{X}{n} - p\right| \leq \frac{1{,}96\sigma}{n}$ erfüllt, d. h. es muss $|0{,}1 - p| \leq \frac{1{,}96\sigma}{n}$ gelten.

Dabei ist $\sigma = \sqrt{n \cdot p \cdot (1 - p)}$ die Standardabweichung der Binomialverteilung zu den Parametern n und p.

Durch Einsetzen von $n = 1352$ und $\sigma = \sqrt{1352 \cdot p \cdot (1 - p)}$ ergibt sich:

$|0{,}1 - p| \leq \frac{1{,}96}{1352} \cdot \sqrt{1352 \cdot p \cdot (1 - p)}$

$(0{,}1 - p)^2 \leq \frac{3{,}8416}{1352} \cdot p \cdot (1 - p) = \frac{3{,}8416}{1352} \cdot (p - p^2)$

$0{,}01 - 0{,}2p + p^2 \leq cp - cp^2 \left(\text{mit } c = \frac{3{,}8416}{1352}\right)$

$(c + 1)p^2 - (c + 0{,}2)p + 0{,}01 \leq 0$

Die Grenzen des gesuchten Konfidenzintervalls sind die Lösungen der quadratischen Gleichung

$(c+1)p^2 - (c+0{,}2)p + 0{,}01 = 0,$

also nach der quadratischen Lösungsformel $\left(\text{mit } c = \frac{3{,}8416}{1352}\right)$:

$$p_1 = \frac{c+0{,}2-\sqrt{(c+0{,}2)^2 - 4\cdot(c+1)\cdot 0{,}01}}{2\cdot(c+1)}$$
$$\approx 0{,}08512$$

und:

$$p_2 = \frac{c+0{,}2+\sqrt{(c+0{,}2)^2 - 4\cdot(c+1)\cdot 0{,}01}}{2\cdot(c+1)}$$
$$\approx 0{,}1171.$$

Somit ist $K_G = [0{,}0851; 0{,}1171]$ näherungsweise ein 95 %-Konfidenzintervall für p_G.

Aufgabe c

1. SCHRITT: 95 %-KONFIDENZINTERVALL FÜR p_G BESTIMMEN

Sei p_F der Anteil der Wähler, die sich am nächsten Sonntag für die FDP entscheiden würden. Sei ferner X die Anzahl der FDP-Wähler unter den 1352 Befragten. X wird als binomialverteilt zu den Parametern $n = 1352$ und p_F angenommen.

Dementsprechend ist der Erwartungswert von X

$\mu = \mathrm{E}(X) = n \cdot p_F = 1352 \cdot p_F$

und sofern die Laplace-Bedingung erfüllt ist, gilt laut Tabelle 1 (σ-Regeln)

$P(\mu - 1{,}96\sigma \le X \le \mu + 1{,}96\sigma) \approx 0{,}95$, d. h.

$$P(|X-\mu| \le 1{,}96\sigma) \approx 0{,}95 \Leftrightarrow P\left(\left|\frac{X}{n} - p_F\right| \le \frac{1{,}96\sigma}{n}\right) \approx 0{,}95.$$

Gesucht ist ein Teilintervall $K_F \subset [0; 1]$, das (in einem geeigneten Sinne) mit mindestens 95 % Wahrscheinlichkeit p_F enthält. Gemeint ist: K_F enthält genau die Schätzwerte $p \in [0; 1]$, für die das Ereignis $\frac{X}{n} \approx 14{,}6$ %, das bei der Wahl eingetreten ist, die Bedingung $\left|\frac{X}{n} - p\right| \le \frac{1{,}96\sigma}{n}$ erfüllt, d. h. es muss $|0{,}146 - p| \le \frac{1{,}96\sigma}{n}$ gelten.

Dabei ist $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ die Standardabweichung der Binomialverteilung zu den Parametern n und p.

Durch Einsetzen von $n = 1352$ und $\sigma = \sqrt{1352 \cdot p \cdot (1-p)}$ ergibt sich:

$$|0{,}146 - p| \le \frac{1{,}96}{1352} \cdot \sqrt{1352 \cdot p \cdot (1-p)}$$

$$(0{,}146 - p)^2 \le \frac{3{,}8416}{1352} \cdot p \cdot (1-p) = \frac{3{,}8416}{1352} \cdot (p - p^2)$$

$0{,}021316 - 0{,}292p + p^2 \leq cp - cp^2 \left(\text{mit } c = \frac{3{,}8416}{1352}\right)$

$(c+1)p^2 - (c+0{,}292)p + 0{,}021316 \leq 0$

Die Grenzen des gesuchten Konfidenzintervalls sind die Lösungen der quadratischen Gleichung

$(c+1)p^2 - (c+0{,}292)p + 0{,}021316 = 0,$

also nach der quadratischen Lösungsformel $\left(\text{mit } c = \frac{3{,}8416}{1352}\right)$:

$$p_1 = \frac{c + 0{,}292 - \sqrt{(c+0{,}292)^2 - 4 \cdot (c+1) \cdot 0{,}021316}}{2 \cdot (c+1)} \approx 0{,}1282$$

und:

$$p_2 = \frac{c + 0{,}292 - \sqrt{(c+0{,}292)^2 - 4 \cdot (c+1) \cdot 0{,}021316}}{2 \cdot (c+1)} \approx 0{,}1658.$$

Somit ist $K_F = [0{,}1282; 0{,}1658]$ näherungsweise ein 95 %-Konfidenzintervall für p_F.

Der Umfragewert von 13 % liegt in diesem Konfidenzintervall, ist also mit 95 %iger Sicherheit verträglich mit dem Wahlergebnis. Insbesondere wird die Behauptung, dass die Kurzentschlossenheit die Wähler die Aussagekraft von Umfragen zunichte mache, nicht bestätigt.

Für kleine Abweichung des Umfrageergebnisses vom Wahlergebnis können durchaus andere Ursachen verantwortlich sein als die Kurzentschlossenheit der Wähler.

Aufgabe d

1. SCHRITT: SICHERHEITSWAHRSCHEINLICHKEIT BERECHNEN

Die Sicherheitswahrscheinlichkeit für den Schätzwert p aus der Umfrage ist

$$P\left(\left|\frac{X}{n} - p\right| \leq \varepsilon\right),$$

wobei ε der Fehlerbereich ist. Für kleine Parteien mit $p \approx 10$ % wird $\varepsilon \approx 0{,}02$ angegeben.

Für $n = 1352$ und $p = 10$ % ist $\mu = n \cdot p = 135{,}2$ und $\sigma = \sqrt{n \cdot p \cdot (p-1)} \approx 11{,}03$. Also ist die Laplace-Bedingung $\sigma > 3$ erfüllt und wir benutzen daher eine Näherung mit der Normalverteilung: Für $z = \frac{\varepsilon n}{\sigma} \approx 2{,}45$ gilt

$$\begin{aligned} P\left(\left|\frac{X}{n}-p\right| \le \varepsilon\right) &= P\left(\left|\frac{X}{n}-p\right| \le \frac{z\sigma}{n}\right) \\ &\approx \Phi(z)-\Phi(-z) \\ &\approx \Phi(2{,}45)-\Phi(-2{,}45) \\ &= \Phi(2{,}45)-\big(1-\Phi(2{,}45)\big) \\ &= 2\Phi(2{,}45)-1 \\ &\approx 2\cdot 0{,}9929-1 \\ &= 0{,}9858. \end{aligned}$$

(Den Wert $\Phi(2{,}45) \approx 0{,}9929$ entnimmt man Tabelle 6.)

2. SCHRITT: VERGLEICH UND INTERPRETATION

Insgesamt ist die Sicherheitswahrscheinlichkeit für die Prognosen bei der Forschungsgruppe Wahlen sehr hoch. Obwohl der Fehlerbereich bei den kleineren Parteien kleiner ist (± 2 % gegenüber ± 3 % bei den größeren), ist die Sicherheitswahrscheinlichkeit bei den kleineren Parteien mit 98,58 % noch größer als bei den größeren Parteien mit 97,56 %.