Abitur Mathematik: Musterlösung

Aufgabe 7: Stochastik (WTR)

Nordrhein-Westfalen 2013

Aufgabe 7

a)

SITUATION MODELLIEREN

Annahmen:

- Es werden 100 Personen unabhängig voneinander befragt.
- Auf die Frage, ob mindestens einmal im Monat ein Fahrrad genutzt wurde, gibt es genau zwei mögliche Antworten: „Ja“ oder „Nein“.
- Die Antwort ist bei jeder Befragung mit Wahrscheinlichkeit $\frac{2}{3}$ „Ja“ und mit Wahrscheinlichkeit $\frac{1}{3}$ „Nein“.

Die Anzahl X der Personen, die mit „Ja“ antworten, ist somit binomialverteilt zu den Parametern $n = 100$ und $p = \frac{2}{3}$.

WAHRSCHEINLICHKEIT FÜR E_1 BERECHNEN

Gesucht ist:

$P(E_1) = P(X = 70)$.

Die Bernoulli-Formel liefert:

$$P(X = 70) = \binom{100}{70} \cdot \left(\frac{2}{3}\right)^{70} \cdot \left(\frac{1}{3}\right)^{30} \approx 0{,}0673.$$

Mit einer Wahrscheinlichkeit von etwa 6,7 % benutzen genau 70 der 100 Befragten ihr Fahrrad mindestens einmal im Monat.

WAHRSCHEINLICHKEIT FÜR E_2 BERECHNEN

Gesucht ist:

$P(E_2) = P(X \geq 70) = 1 - P(X \leq 69)$.

In Tabelle 4 (kumulierte Binomialverteilung für $n = 100$) wird $1 - P(X \leq 69)$ direkt angegeben: $p = \frac{2}{3}$ ist in der unteren grau eingefärbten

Zeile aufgelistet und in der rechten grau eingefärbten Spalte der Wert $k = 69$. Es ist also:

$1 - P(X \leq 69) \approx 0{,}2766$.

Mit einer Wahrscheinlichkeit von ca. 27,7 % benutzen mindestens 70 von den 100 Befragten ihr Fahrrad mindestens einmal im Monat.

WAHRSCHEINLICHKEIT FÜR E_3 BERECHNEN

Gesucht ist:

$P(E_3) = P(60 \leq X \leq 70) = P(30 \leq Y \leq 40) = P(Y \leq 40) - P(Y \leq 29)$,

mit Y = Anzahl der Personen, die mit „Nein" antworten; für diese Zufallsgröße ist $p = \frac{1}{3}$.

Aus Tabelle 4 können die Werte $P(Y \leq 40) \approx 0{,}9341$ und $P(Y \leq 29) \approx 0{,}2093$ entnommen werden. Demzufolge ist:

$P(Y \leq 40) - P(Y \leq 29) = P(60 \leq X \leq 70)$

$\approx 0{,}9341 - 0{,}2093 = 0{,}7248$.

Mit einer Wahrscheinlichkeit von ca. 72,5 % benutzen mindestens 60 und höchstens 70 von 100 Befragten ihr Fahrrad mindestens einmal im Monat.

b)

SITUATION MODELLIEREN

Annahmen:

- Es werden n Fahrräder unabhängig voneinander kontrolliert.
- Für jede Fahrradkontrolle gibt es genau zwei mögliche Ausgänge: „Beanstandung" oder „keine Beanstandung".
- Die Wahrscheinlichkeit einer Beanstandung ist bei jeder Kontrolle $\frac{1}{6}$.

Die Anzahl X der Beanstandungen ist somit binomialverteilt zu den Parametern n und $p = \frac{1}{6}$.

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

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

$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{n}{0} \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^n = 1 - \left(\frac{5}{6}\right)^n$.

Die zu erfüllende Ungleichung lautet also

$1 - \left(\frac{5}{6}\right)^n \geq 0{,}9 \qquad \left| -0{,}9 + \left(\frac{5}{6}\right)^n \right.$

$\left(\frac{5}{6}\right)^n \leq 0{,}1 \qquad |$ logarithmieren

$\ln\left(\left(\frac{5}{6}\right)^n\right) \le \ln 0{,}1$ | Logarithmengesetz anwenden

$n \cdot \ln\left(\frac{5}{6}\right) \le \ln 0{,}1$ $\Big| : \ln\left(\frac{5}{6}\right)$

$n \ge \frac{\ln 0{,}1}{\ln\left(\frac{5}{6}\right)} \approx 12{,}63$

Die kleinste natürliche Zahl mit dieser Eigenschaft ist 13.

Um mit 90 %iger Wahrscheinlichkeit mindestens ein Fahrrad mit Mängeln zu finden, müssen mindestens 13 Fahrräder kontrolliert werden.

c)

(1) WAHRSCHEINLICHKEIT FÜR GROẞSTADTHERKUNFT UND FAHRRADNUTZUNG BESTIMMEN

BEZEICHNUNGEN:

K = Person wohnt in Ort mit unter 20.000 Einwohnern
M = Person wohnt in Ort mit 20.000 bis 100.000 Einwohnern
G = Person wohnt in Ort mit über 100.000 Einwohnern
R = Person nutzt regelmäßig ein Fahrrad
$\overline{R}$ = Person nutzt nicht regelmäßig ein Fahrrad

Gesucht ist:

$P(G \cap R) = P(G) \cdot P_G(R)$

Aus der Tabelle entnimmt man: $P(G) = 0{,}306$; $P_G(R) = 0{,}43$. Damit ist:

$P(G \cap R) = 0{,}306 \cdot 0{,}43 = 0{,}13158$

Mit einer Wahrscheinlichkeit von etwa 13 % kommt eine Person aus einer Stadt mit mehr als 100000 Einwohnern und benutzt regelmäßig ihr Fahrrad.

(2) WAHRSCHEINLICHKEIT FÜR GROẞSTADTHERKUNFT UNTER DER VORAUSSETZUNG NICHT REGELMÄẞIGER FAHRRADNUTZUNG

Gesucht ist:

$P_{\bar{R}}(G) = \frac{P(G \cap \bar{R})}{P(\bar{R})}$

Es gilt: $P_G(\bar{R}) = 1 - P_G(R) = 1 - 0{,}43 = 0{,}57$

Damit folgt:

$P(G \cap \bar{R}) = P(G) \cdot P_G(\bar{R}) = 0{,}306 \cdot 0{,}57 = 0{,}17442.$

Analog ist:

$P(K \cap \overline{R}) = P(K) \cdot P_K(\overline{R}) = 0{,}404 \cdot (1 - 0{,}37) = 0{,}25452,$

$P(M \cap \overline{R}) = P(M) \cdot P_M(\overline{R}) = 0{,}29 \cdot (1 - 0{,}42) = 0{,}1682.$

Nach der 2. Pfadregel ist

$$\begin{aligned} P(\bar{R}) &= P(K \cap \bar{R}) + P(M \cap \bar{R}) + P(G \cap \bar{R}) \\ &= 0{,}25452 + 0{,}1682 + 0{,}17442 \\ &= 0{,}5971 \end{aligned}$$

Damit ist

$$P_{\bar{R}}(G) = \frac{0{,}17442}{0{,}5971} = 0{,}2921.$$

Mit einer Wahrscheinlichkeit von 29,21 % kommt eine Person aus einer Großstadt, wenn sie nicht regelmäßig ein Fahrrad benutzt.

d)

MINDESTANTRITTSGEBÜHR BESTIMMEN

ANNAHMEN

Aus der Erfahrung des Vorjahres wird extrapoliert, dass 25 % der potentiellen Teilnehmer für das aktuelle Jahr das Rennen in der vorgegebenen Zeit schaffen. Die Zusammenstellung der 200 Teilnehmer für das diesjährige Rennen entspricht dann dem Modell „Ziehen ohne Zurücklegen“ aus der Menge der potentiellen Teilnehmer. Da aber die Zahl der potentiellen Teilnehmer vermutlich weitaus größer ist, als 200, kann näherungsweise das Modell „Ziehen ohne Zurücklegen“ verwendet werden, d.h. die Anzahl X der Radfahrer, die einen Einkaufsgutschein bekommen, ist binomialverteilt zu den Parametern $n = 200$ und $p = 0{,}25$.

ABSCHÄTZUNG DER ANZAHL DER GEWINNER

Der Veranstalter muss mit 95 %iger Sicherheit abschätzen, wie viele Teilnehmer höchstens einen Einkaufsgutschein gewinnen. Er sucht also das kleinste k, so dass

$P(X \leq k) \geq 0{,}95$ ist.

Da X als binomialverteilt angenommen wird, hat X die Standardabweichung

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)} = \sqrt{200 \cdot 0{,}25 \cdot 0{,}75} \approx 6{,}12 > 3.$$

Somit ist die Laplace-Bedingung erfüllt und es gilt daher nach den σ-Regeln

$P(X \leq \mu + 1{,}64\sigma) \approx 0{,}95,$

wobei $\mu = n \cdot p = 200 \cdot 0{,}25 = 50$ der Erwartungswert von X ist.

Somit ist $\mu + 1{,}64\sigma \approx 60{,}04$. Um die 95 %ige Sicherheit nicht zu gefährden wird der nächstgrößere ganzzahlige Wert verwendet: $k = 61$.

Der Veranstalter kann mit 95 %iger Sicherheit damit rechnen, dass höchstens 61 Fahrradfahrer einen Preis gewinnen.

MINDESTSTARTGEBÜHR

Jeder der höchstens 61 Gewinner kostet dem Veranstalter 20€, d.h. er rechnet mit Ausgaben von höchstens $61 \cdot 20€ = 1220€$.

Die Einnahmen durch die Startgebühr belaufen sich auf

$200 \cdot x$, wobei x die Startgebühr ist.

Es muss also gewährleistet sein, dass $200 \cdot x > 1220€$ ist, d.h. der Veranstalter muss mehr als

$\frac{1220€}{200} = 6{,}10€$ verlangen.

e)

(1) HYPOTHESEN BESTIMMEN UND BEGRÜNDEN

BESTIMMEN DER HYPOTHESEN

Es bezeichne p den Anteil der Fahrräder mit Mängel.

Mögliche Nullhypothesen sind:

entweder

a) $p < 0{,}1$ („Weniger als 10 % der Fahrräder weisen Mängel auf.“)

oder

b) $p \geq 0{,}1$ („Mindestens 10 % der Fahrräder weisen Mängel auf“).

Die Anzahl X der bei der morgendlichen Kontrolle beanstandeten Fahrräder ist binomialverteilt mit Parametern $n = 200$ und unbekanntem p.

Aus Tabelle 5 (kumulierte Binomialverteilung für $n = 200$) entnimmt man für $p = 0{,}1$ die kumulierten Wahrscheinlichkeiten:

$P(X \leq 12) \approx 0{,}032$ sowie $P(X \leq 13) \approx 0{,}0566$.

Das Signifikanzniveau ist die obere Schranke für die Wahrscheinlichkeit, die Nullhypothese irrtümlich zu verwerfen ($\alpha = 0{,}05$).

$P(X \leq 12)$ liegt gerade noch unterhalb dieser Schranke, die (noch unbekannte) Nullhypothese wird demnach für $X \leq 12$ verworfen.

Das Ergebnis $X \leq 12$ deuten darauf hin, dass der Anteil der zu beanstandenden Fahrräder klein ist, dass also $p < 0{,}1$ ist. Wenn in diesem Fall die Nullhypothese verworfen wird, dann muss sie lauten:

Nullhypothese H_0: $p \geq 0{,}1$

Gegenhypothese H_1: $p < 0{,}1$.

BEGRÜNDEN DER HYPOTHESENWAHL

Das Signifikanzniveau ist die obere Schranke für den Fehler 1. Art (Wahrscheinlichkeit, die Nullhypothese irrtümlich zu verwerfen). Die Folgen eines Fehlers 1. Art sind entweder

a) Es wird davon ausgegangen, dass mindestens 10 % der Fahrräder Mängel aufweisen, obwohl in Wirklichkeit weniger als 10 % der Fahrräder zu beanstanden sind, d.h., die Kontrollen werden nicht reduziert, obwohl dies gerechtfertigt wäre

oder

b) Es wird davon ausgegangen, dass höchstens 10 % der Fahrräder Mängel aufweisen, obwohl in Wirklichkeit mehr als 10 % der Fahrräder zu beanstanden sind, d.h., die Kontrollen werden ungerechtfertigterweise reduziert.

Im Fall a) würde die Polizei unnötigerweise auf Ressourcenersparnis verzichten. Im Fall b) wäre die Verkehrssicherheit gefährdet. Es ist also aus Sicht der Polizei sinnvoller, die Wahrscheinlichkeit des Fehlers 1. Art im Fall b) gering zu halten, weswegen $p \geq 0{,}1$ als Nullhypothese gewählt wurde.

(2) WAHRSCHEINLICHKEIT FÜR EINEN FEHLER 1. ART

Im „verdoppelten“ Test erfolgt nach dem Beschluss der Polizisten ein Fehler 1. Art, wenn höchstens 25 defekte Fahrräder gefunden werden, obwohl tatsächlich mindestens 10 % der Räder defekt sind. Die Wahrscheinlichkeit für einen Fehler 1. Art ergibt sich also als $P(Y \leq 25)$, wenn Y die Anzahl der beanstandeten Fahrräder ist. Y ist binomialverteilt mit Parametern $n = 400$ und unbekanntem p.

Aus Tabelle 6 (kumulierte Binomialverteilung für $n = 400$, $p = 0{,}1$) entnimmt man den Wert:

$$P(Y \leq 25) \approx 0{,}0054 \Longrightarrow \frac{P(Y \leq 25)}{P(X \leq 12)} \approx \frac{0{,}0054}{0{,}032} = 0{,}16875 < 0{,}2.$$

Damit ist die Wahrscheinlichkeit, in dem neuen Test einen Fehler 1. Art zu begehen, auf weniger als ein Fünftel der ursprünglichen Fehlerwahrscheinlichkeit gesunken.

(3) ZU ZEIGEN: $\sigma_{400} = \sqrt{2} \cdot \sigma_{200}$

Die Standardabweichung einer binomialverteilten Zufallsgröße mit Parametern n und p berechnet sich gemäß der Formel

$$\sigma = \sqrt{n \cdot p \cdot (1-p)}.$$

Demnach ist

$$\sigma_{200} = \sqrt{200 \cdot p \cdot (1-p)} = \sqrt{200} \cdot \sqrt{p \cdot (1-p)}$$

und

$$\begin{aligned}\sigma_{400} &= \sqrt{400 \cdot p \cdot (1-p)} \\ &= \sqrt{400} \cdot \sqrt{p \cdot (1-p)} \\ &= \sqrt{2 \cdot 200} \cdot \sqrt{p \cdot (1-p)} \\ &= \sqrt{2} \cdot \left(\sqrt{200} \cdot \sqrt{p \cdot (1-p)}\right) \\ &= \sqrt{2} \cdot \sigma_{200}\end{aligned}$$

AUSWIRKUNG AUF DEN ANNAHMEBEREICH ERLÄUTERN

Der Annahmebereich der Nullhypothese H_0: $p \geq 0{,}1$ bei einer Stichprobe der Länge n ist die Menge aller möglichen Ergebnisse k des Experiments X_n, die zur Annahme von H_0 führen. Dieser Bereich hat die Form

$\{k_0 + 1; k_0 + 2; \dots; n\}$ für ein $k_0 \in \mathbb{N}_0$.

Der kritische Wert k_0 wird näherungsweise mit den σ-Regeln bestimmt. Um nämlich das Signifikanzniveau von $\alpha = 0{,}05$ zu halten muss die Anzahl X_n der beanstandeten Fahrräder mindestens mit Wahrscheinlichkeit $1 - \alpha = 0{,}95$ im Annahmebereich liegen, d.h. es muss $P(X_n \geq k_0 + 1) \geq 0{,}95$ gelten.

Nun gilt nach den σ-Regeln

$P(X_n \geq \mu_n - 1{,}64\sigma_n) \approx 0{,}95$, d.h. k_0 wird so gewählt, dass $k_0 + 1$ die größte ganze Zahl mit $k_0 + 1 \leq \mu_n - 1{,}64\sigma_n$ ist. Notation: $k_0 + 1 = \lfloor \mu_n - 1{,}64\sigma_n \rfloor$.

Durch den Übergang von $n = 200$ auf $n = 400$ wird also unter Beibehaltung des Signifikanzniveaus aus dem Annahmebereich

$\{\lfloor \mu_{200} - 1{,}64\sigma_{200} \rfloor; \dots; 200\}$

der Annahmebereich

$\{\lfloor \mu_{400} - 1{,}64\sigma_{400} \rfloor; \dots; 200\} = \{\lfloor 2\mu_{200} - 1{,}64\sqrt{2}\sigma_{200} \rfloor; \dots; 200\}$.

Die linke Grenze des Annahmebereichs wird also nicht nur zu $2k_0 + 2$ verdoppelt, sondern es wird noch $1{,}64(2 - \sqrt{2})\sigma_{200}$ addiert. Der Annahmebereich fällt somit kleiner aus, als bei der Strategie der Polizei. Durch die Wahl eines für das vorgegebene Signifikanzniveau unnötig großen Annahmebereichs ist die Wahrscheinlichkeit des Fehlers 2. Art gestiegen, d.h. es wird möglicherweise keine Reduktion der Kontrollen vorgenommen, obwohl der Anteil mangelhafter Fahrräder unter 10 % gesunken ist.