ler Schule, Heft 2, Band 2 (1982)

Eine Aufgabe zur Stochastik der Sekundarstufe II

G. Fillbrunn

Die folgende Aufgabe eignet sich für die schriftliche Abiturprüfung eines Leistungskurses. Die Bearbeitungszeit beträgt 120 Minuten.

Es soll geschätzt werden, wie hoch unter den Studenten der Anteil p derjenigen ist, die bereits Rauschgift genommen haben. Dazu wird ein Verfahren gewählt, das für die Betroffenen unverfänglich ist und die Intimsphäre schützt. Bei diesem macht jeder Befragte einen Wurf mit einem L-Würfel. Den Ausgang darf nur der Befragte sehen. Ist die Augenzahl 1 oder 2 oder 3, dann soll der Student die Frage 'Haben Sie jemals Rauschgift genommen?' ehrlich beantworten. Bei den Augenzahlen 4 und 5 soll er die Frage mit 'ja' und bei der Augenzahl 6 mit 'nein' beantworten.

a) a_1) Zeichne ein Baumdiagramm für das zweistufige Zufallsexperiment:

1. Teilexperiment: 'Auswahl eines Studenten' [Ausgänge: 'mit Rauschgiftkontakt' (=mR), 'ohne Rauschgiftkontakt' (=oR)]

2. Teilexperiment: 'Würfeln' [Ausgänge: 'Augenzahl 1 oder 2 oder 3' (=123), 'Augenzahl 4 oder 5' (=45) und 'Augenzahl 6' (=6)]

a_2) Gib für das Ereignis, daß ein Student befragt wird, der eine Antwort gibt, die nicht der Wirklichkeit hinsichtlich des Rauschgiftkontakts entspricht, die Wahrscheinlichkeit in Abhängigkeit von p an.

a_3) Gib für das Ereignis, daß ein Student befragt wird, der mit 'ja' antwortet, die Wahrscheinlichkeit in Abhängigkeit von p an.

b) Aus einer sehr großen Studentenpopulation werden 1000 zufällig gewählte Personen befragt. Die Zufallsvariable X gibt an, wieviel Personen mit 'ja' antworten.

b_1) Man gebe die Verteilung von X an.

b_2) Berechne E(X) und V(X).

b_3) 382 der 1000 befragten Personen antworten mit 'ja'. Berechne einen Schätzwert für p.

c) Wir gehen nun davon aus, daß $p = 0{,}10$ gilt. Gib für die Wahrscheinlichkeit, daß sich in einer Stichprobe vom Umfang

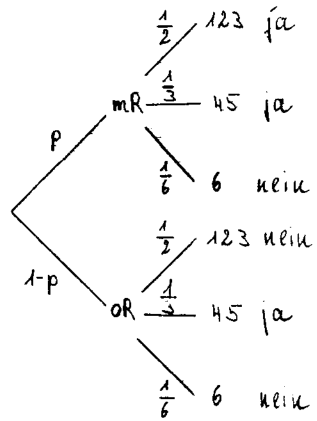
c_1) 20 mindestens 3 Studenten

c_2) 1000 mindestens 105 Studenten

befinden, die bereits Rauschgift zu sich genommen haben, einen gerundeten Wert an.

d) Unter den Studentinnen ist p_w der Anteil derjenigen, die bereits Rauschgift genommen haben. Um p_w zu schätzen, werden 1000 Studentinnen nach dem eingangs geschilderten Verfahren befragt. 368 Personen antworten mit 'ja'.
Ermittle mit Hilfe der Normalverteilungsapproximation der Binomialverteilung ein 0,90-Vertrauensintervall für p_w.

Lösung: a) a_1)



a_2) Eine falsche Antwort erfolgt bei den Ausgängen (mR,6) und (oR,45). Also

$$p_{a_2} = \frac{1}{6}p + \frac{1}{3}(1-p)$$
$$= \frac{1}{3} - \frac{1}{6}p$$

a_3) Die Antwort 'ja' erhält man bei (mR,123), (mR,45), (oR,45).

$$p_{a_3} = \frac{1}{2}p + \frac{1}{3}p + \frac{1}{3}(1-p)$$
$$= \frac{1}{2}p + \frac{1}{3}$$

b) b_1) $$k \mapsto \binom{1000}{k}\cdot\left(\frac{1}{2}p+\frac{1}{3}\right)^k\cdot\left(\frac{2}{3}-\frac{1}{2}p\right)^{1000-k}, \quad k \in \{0,\ldots,1000\}$$

b_2) $$E(X) = 1000\cdot\left(\frac{1}{2}p+\frac{1}{3}\right)$$
$$V(X) = 1000\cdot\left(\frac{1}{2}p+\frac{1}{3}\right)\cdot\left(\frac{2}{3}-\frac{1}{2}p\right)$$

b_3) Der Schätzwert $\overline{y}$ von p wird aus

$$382 = 1000\cdot\left(\frac{1}{2}\overline{y} + \frac{1}{3}\right)$$

ermittelt:

$$\frac{1}{2}\bar{y}+\frac{1}{3} = \frac{382}{1000}$$

$$\bar{y} = \frac{764}{1000}-\frac{2}{3}$$

$$\approx 0{,}0973$$

c) c_1)

$$p_{c_1} = \sum_{k=3}^{20} B_{20;0,1}(k)$$

$$= \sum_{k=3}^{20}\binom{20}{k}\cdot 0{,}1^k\cdot 0{,}9^{20-k}$$

$$\approx 1-0{,}6769 = 0{,}3231$$

c_2)

$$p_{c_2} = \sum_{k=105}^{1000} B_{1000;0,1}(k)$$

$$= \sum_{k=105}^{1000}\binom{1000}{k}\cdot 0{,}1^k\cdot 0{,}9^{1000-k}$$

Da die $B_{1000;0,1}$-verteilte Zufallsvariable näherungsweise $N(100;\sqrt{90})$-verteilt ist, wird ein gerundeter Wert von p_{c_2} wie folgt gewonnen:

$$p_{c_2} \approx 1-\Phi\left(\frac{104{,}5-100}{\sqrt{90}}\right)$$

$$\approx 1-\Phi(0{,}4743)$$

$$\approx 1-0{,}6824 \approx 0{,}32$$

d) Die Wahrscheinlichkeit, daß eine Studentin eine Ja-Antwort gibt, ist nach a_3) $\bar{p}$ mit $\bar{p}=\frac{1}{2}p_w+\frac{1}{3}$.
Sind $\bar{Z}, S$ die Zufallsvariablen für die relative Häufigkeit und für die empirische Standardabweichung der Ja-Antworten bei 1000 befragten Studentinnen, darf

$$\sqrt{1000}\cdot\frac{\bar{Z}-(\frac{1}{2}p_w+\frac{1}{3})}{S}$$

als näherungsweise $N(0;1)$-verteilt angesehen werden.
Nun ist die in der Klammer von

$$P\left(-1{,}645\le\sqrt{1000}\cdot\frac{\bar{Z}-(\frac{1}{2}p_w+\frac{1}{3})}{S}\le 1{,}645\right)\approx 0{,}90$$

stehende Ungleichung umzuformen:

$$\bar{Z}-1{,}645\cdot\frac{S}{\sqrt{1000}}\le\frac{1}{2}p_w+\frac{1}{3}\le\bar{Z}+1{,}645\cdot\frac{S}{\sqrt{1000}}$$

$$2\bar{Z}-3{,}290\cdot\frac{S}{\sqrt{1000}}-\frac{2}{3}\le p_w\le 2\bar{Z}+3{,}290\cdot\frac{S}{\sqrt{1000}}-\frac{2}{3}$$

Durch Realisierung wird dann

$$\bar{z} = 0{,}368$$

$$s^2 = \frac{1}{999}\cdot\left(368-\frac{1}{1000}\cdot 368^2\right)$$

$$= \frac{368}{999}\cdot(1-0{,}368)$$

$$\approx 0{,}2328$$

$$s \approx 0{,}4825$$

und damit das Vertrauensintervall

$$\left[0{,}736-3{,}290\cdot\frac{0{,}4825}{\sqrt{1000}}-\frac{2}{3}\,;\,0{,}736+3{,}290\cdot\frac{0{,}4825}{\sqrt{1000}}-\frac{2}{3}\right],$$

also

$$[0{,}0191;0{,}1195]$$

gewonnen.

Anmerkung: Bei Teil d) könnte ein Schüler auf folgende 'Lösung' verfallen:
Mit dem Schätzwert $\frac{368}{1000}$ für $\bar{p}$ wird wie in b_3) der Schätzwert $2\cdot(\frac{368}{1000}-\frac{1}{3})$ ($\approx 0{,}0693$) für p_w berechnet.
Mit den Zufallsvariablen $\tilde{Z}, \tilde{S}$ für die relative Häufigkeit und die empirische Standardabweichung der mit Rauschgift in Kontakt geratenen Studentinnen wird dann

$$\sqrt{1000}\cdot\frac{\tilde{Z}-p_w}{\tilde{S}}$$

als ungefähr $N(0;1)$-verteilt angenommen.
Aus

$$P\left(-1{,}645\le\sqrt{1000}\cdot\frac{\tilde{Z}-p_w}{\tilde{S}}\le 1{,}645\right)\approx 0{,}90$$

erhält man

$$P\left(\tilde{Z}-1{,}645\cdot\frac{\tilde{S}}{\sqrt{1000}}\le p_w\le\tilde{Z}+1{,}645\cdot\frac{\tilde{S}}{\sqrt{1000}}\right)\approx 0{,}90$$

und daraus mit den Realisierungen

$$\tilde{z}\approx 0{,}0693\quad,\quad\tilde{s}\approx\sqrt{0{,}0693\cdot(1-0{,}0693)}\approx 0{,}2540$$

das Vertrauensintervall

$$\left[0{,}0693-1{,}645\cdot\frac{0{,}2540}{\sqrt{1000}}\,;\,0{,}0693+1{,}645\cdot\frac{0{,}2540}{\sqrt{1000}}\right],$$

also

$$[0{,}0561;0{,}0825]\quad.$$

Das so erhaltene Intervall ist viel kürzer als das richtige Vertrauensintervall. Der Fehler liegt in der Annahme, daß in der Stichprobe der tatsächliche Anteil der Studentinnen mit Rauschgiftkontakt 6,93% ist. Die spezielle Art der Befragung läßt diesen Schluß aber nicht zu.
Der Vorteil, daß die Fragestellung unverfänglich ist, muß also mit einer Vergrößerung des Vertrauensintervalls erkauft werden.

Literatur: G. Pflug: 'Paradoxien der Wahrscheinlichkeitsrechnung'
in STOCHASTIK IM SCHULUNTERRICHT , Beiträge zum
3. internationalen Symposium für 'Didaktik der Mathematik'
Hölder-Pichler-Tempsky, Wien
B.G. Teubner, Stuttgart 1981