

6 Serie — Aufgaben zur Statistik

Aufgabe 1.

Erzeugen Sie mit SPSS und R 1000 hypergeometrisch verteilte Zufallszahlen für die Anzahl der weißen Kugeln bei 30 Ziehungen (ohne Zurücklegen) aus einer Urne mit 100 weißen und 200 schwarzen Kugeln.

a) Bestimmen Sie mit SPSS und R und theoretisch die Einzelwahrscheinlichkeiten, dass genau k ($k = 1, 2, \dots, 30$) weiße Kugeln gezogen werden.

b) Untersuchen Sie das monotone Verhalten der Einzelwahrscheinlichkeiten theoretisch und mit dem Programm SPSS und R.

c) Wiederholen Sie das beschriebene Experiment leicht modifiziert: Erzeugen Sie mit SPSS und R 1000 binomialverteilte Zufallszahlen für die Anzahl der weißen Kugeln bei 30 Ziehungen (mit Zurücklegen) aus einer Urne mit 100 weißen und 200 schwarzen Kugeln. Bestimmen Sie die entsprechenden Einzelwahrscheinlichkeiten und untersuchen Sie das monotone Verhalten dieser Wahrscheinlichkeiten theoretisch und mit SPSS und R.

Lösung 1. a) Es sei $l = 30$, $w = 100$ und $s = 200$. Weiter sei $n = s + w$. Dann gilt

$$p_k = \frac{\binom{w}{k} \binom{s}{l-k}}{\binom{n}{l}}.$$

Wir betrachten

$$\begin{aligned} \frac{p_k}{p_{k-1}} &= \frac{\binom{w}{k} \binom{s}{l-k}}{\binom{w}{k-1} \binom{s}{l+1-k}} \\ &= \frac{w!s!(k-1)!(w+1-k)!(l+1-k)!(s+k-l-1)!}{k!(w-k)!(l-k)!(s+k-l)!w!s!} \\ &= \frac{(w+1-k)(l+1-k)}{k(s+k-l)} \end{aligned}$$

Die Ungleichung

$$\frac{p_k}{p_{k-1}} = \frac{(w+1-k)(l+1-k)}{k(s+k-l)} > 1$$

ist gleichwertig mit

$$(w+1-k)(l+1-k) > k(s+k-l)$$

und somit mit

$$(w+1)(l+1) - (w+l+2)k + k^2 > k(s-l) + k^2.$$

Dies ist äquivalent zu

$$(w+1)(l+1) > (n+2)k.$$

also

$$a := \frac{(w+1)(l+1)}{n+2} > k.$$

Also ist die Folge $(p_k)_{k=1,2,\dots}$ zunächst monoton wachsend und dann monoton fallend.

(i) Es sei a nicht ganz und sei k_0 die größte ganze Zahl, für die $k_0 < a$ gilt, so folgt

$$p_1 < p_2 < \dots < p_{k_0} > p_{k_0+1} > \dots$$

(ii) Es sei a ganz, so folgt

$$p_1 < p_2 < \dots < p_{a-1} = p_a > p_{a+1} > \dots$$

In unserem Beispiel haben wir

$$a = \frac{101}{302} 31 = 10,36$$

Also trifft der Fall (i) zu und es gilt

$$p_1 < p_2 < \dots < p_{10} > p_{11} > \dots > p_{30}$$

b) Wir haben für $n = 30$, $p = \frac{1}{3}$ und $q = 1-p$

$$p_j = \binom{n}{j} p^j q^{n-j}$$

Weiter betrachten wir

$$\frac{p_j}{p_{j-1}} = \frac{\binom{n}{j} p}{\binom{n}{j-1} q} = \frac{(n+1-j)p}{jq}$$

Also ist

$$\frac{p_j}{p_{j-1}} = \frac{(n+1-j)p}{jq} > 1$$

gleichwertig mit

$$a := (n+1)p > j$$

Also ist die Folge $(p_j)_{j=1,2,\dots,n}$ zunächst monoton wachsend und dann monoton fallend.

(i) Es sei b nicht ganz und sei j_0 die größte ganze Zahl, für die $j_0 < b$ gilt, so folgt

$$p_1 < p_2 < \dots < p_{j_0} > p_{j_0+1} > \dots > p_n$$

(i i) Es sei a ganz, so folgt

$$p_1 < p_2 < \dots < p_{b-1} = p_b > p_{b+1} > \dots > p_n$$

In unserem Beispiel haben wir

$$b = \frac{31}{3} = 10,33..$$

Also trifft der Fall (i) zu und es gilt

$$p_1 < p_2 < \dots < p_{10} > p_{11} > \dots > p_{30}$$

Aufgabe 2.

Die Verteilung der Blutgruppen wird durch folgende Tabelle beschrieben

Blutgruppe	A	B	AB	0
Wahrscheinlichkeit	0,425	0,145	0,065	0,365

Es werden von 12 Personen die Blutgruppe untersucht. Es sei X_A die Anzahl der Personen mit Blutgruppe A. Analog werden die Zufallsgrößen X_B , X_{AB} und X_0 definiert.

a) Geben Sie die Verteilung dieser Zufallsgrößen an.

b) Wie groß ist Wahrscheinlichkeit(p_1), dass 8 Personen die Blutgruppe A, 4 Personen die Blutgruppe B haben.

c) Mit welcher Wahrscheinlichkeit(p_2) haben unter 12 willkürlich herausgegriffenen Personen 5 bis 7 die Blutgruppe A.

d) Lösen Sie die Aufgabe mit Hilfe von SPSS und R

Lösung 2. a) Der Zufallsvektor (X_A, X_B, X_{AB}, X_0) ist polynomialverteilt. Also gilt

$$\begin{aligned} P(X_A = x_A, X_B = x_B, X_{AB} = x_{AB}, X_0 = x_0) \\ = \frac{12!}{x_A! x_B! x_{AB}! x_0!} p_A^{x_A} p_B^{x_B} p_{AB}^{x_{AB}} p_0^{x_0} \end{aligned}$$

b)

$$P(X_A = 8, X_B = 4) = \frac{12!}{8!4!} p_A^8 p_B^4 = \frac{2,329}{10^4}$$

c) Es gilt $X_A \sim \text{Bin}(n, p_A)$. Es sei $q_A = 1 - p_A$. Dann folgt

$$\begin{aligned} P(5 \leq X_A \leq 7) &= \binom{12}{5} p_A^5 q_A^7 + \binom{12}{6} p_A^6 q_A^6 + \binom{12}{7} p_A^7 q_A^5 \\ &= 0,2282 + 0,1968 + 0,12347 = 0,5497 \end{aligned}$$

Aufgabe 3.

Für $i = 1, 2, \dots, 1000$ seien die Zufallsgrößen $X_i \sim N(0, 1)$ und $Y_i \sim N(3, 4)$ unabhängig.

- a) Erzeugen Sie mit SPSS und R solch eine Stichprobe.
- b) Erzeugen Sie daraus eine Stichprobe für die abgeleiteten Merkmale $Z_i = 3X_i + 5Y_i$ und $W_i = Y_i/X_i$.
- c) Zeichnen Sie für die Merkmale Z und W die Histogramme.
- d) Welche Verteilung hat Z ? Bestimmen Sie EZ und $\text{Var}(Z)$.
- e) Wie kann man aus der Stichprobe die Parameter der Verteilung von Z , wie Erwartung und Varianz schätzen?
- f) Bestimmen Sie die bedingte Verteilung von W unter $X=x$. Zeigen Sie, dass die Zufallsgröße W keine Erwartung besitzt. Bestimmen Sie die Verteilungsfunktion von W .

Lösung 3. d) Die Zufallsgröße Z ist normalverteilt mit dem Erwartungswert

$$EZ_i = 3EX_i + 5EY_i = 15$$

und der Varianz

$$\text{Var}(Z_i) = 3^2 \text{Var}(X_i) + 5^2 \text{Var}(Y_i) = 109$$

Also gilt $Z_i \sim N(15, 109)$.

d) Eine Schätzung für EZ_i ist

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$$

und eine Schätzung für $\text{Var}(Z_i)$ ist

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2.$$

f) Es gilt

$$P(W_i \leq w | X_i = x) = P\left(\frac{Y_i}{x} \leq w | X_i = x\right) = P\left(\frac{Y_i}{x} \leq w\right)$$

Also folgt als Verteilung

$$N\left(\frac{3}{x}, \frac{4}{x^2}\right)$$

Da das Integral

$$\int_{(0, \varepsilon)} \frac{1}{x} e^{-\frac{x^2}{2}} dx$$

nicht existiert existiert auch der Erwartungswert von W nicht. Für die Verteilungsfunktion F von W erhalten wir, da $X \sim -X$

$$\begin{aligned}
 F(w) &= P\left(\frac{Y}{X} \leq w\right) \\
 &= P\left(\frac{Y}{X} \leq w, X > 0\right) + P\left(\frac{Y}{X} \leq w, X < 0\right) \\
 &= P\left(\frac{Y}{X} \leq w, X > 0\right) + P\left(\frac{Y}{-X} \leq -w, -X > 0\right) \\
 &= P\left(\frac{Y}{X} \leq w, X > 0\right) + P\left(\frac{Y}{X} \leq -w, X > 0\right) \\
 &= P(Y \leq wX, X > 0) + P(Y \leq -wX, X > 0) \\
 &= \int_0^\infty \left(\Phi\left(\frac{xw-3}{4}\right) + \Phi\left(\frac{xw+3}{4}\right) \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.
 \end{aligned}$$