Übungen zur Wahrscheinlichkeitsrechnung und Statistik für Biologen

Blatt 3

---

1. Aufgabe (Berechnung von Erwartungswerten) Sei X eine Zufallsvariable mit möglichen Werten in $\{1, 2, 3\}$ und Verteilung

$$\Pr(X = 1) = \frac{1}{2} \quad \Pr(X = 2) = \frac{1}{3} \quad \Pr(X = 3) = \frac{1}{6}$$

Berechnen Sie per Hand

(a) den Erwartungswert $\mathbb{E}X$ von X (Ergebnis: $\frac{5}{3}$),

(b) $\mathbb{E}(X^2)$,

(c) die Varianz $\mathrm{Var}(X)$ von X,

(d) $\mathbb{E}(X^2 + X)$, und

(e) $\mathbb{E}\Big[(X + 1)^2\Big]$.

(Hinweis: Verwenden Sie dazu die Definition des Erwartungswerts und/oder Linearität.)

2. Aufgabe Sei X die Zufallsvariable aus Aufgabe 1. Desweiteren sei Y eine Zufallsvariable mit Werten in $\{0, 1\}$. Die gemeinsame Verteilung von (X, Y) sei

$$\begin{aligned}
\Pr(X = 1, Y = 0) &= \frac{1}{2} \qquad \Pr(X = 1, Y = 1) = 0 \\
\Pr(X = 2, Y = 0) &= \frac{1}{6} = \Pr(X = 2, Y = 1) \\
\Pr(X = 3, Y = 0) &= \frac{1}{12} = \Pr(X = 3, Y = 1).
\end{aligned}$$

Berechnen Sie per Hand

(a) $\Pr(Y = 0)$ und $\Pr(Y = 1)$,

(b) den Erwartungswert $\mathbb{E}Y$ von Y,

(c) $\mathbb{E}(Y^2)$,

(d) die Standardabweichung $\sqrt{\mathrm{Var}(Y)}$ von Y,

(e) die Kovarianz $\mathrm{Cov}(X, Y)$ von X und Y, und

(f) die Korrelation $\mathrm{Cor}(X, Y)$ zwischen X und Y.

3. Aufgabe Sei $n \in \mathbb{N}$ und seien $x_1, x_2, \ldots, x_n$ reelle Zahlen (z.B. $n = 5$ und $x_1 = 2.5$, $x_2 = 5$, $x_3 = 5$, $x_4 = -1$, $x_5 = -1$). Ziehe rein zufällig eine dieser Zahlen. Sei X die gezogene Zahl. (Zur besseren Vorstellung: Beschrifte n Kugeln gleicher Bauart mit den Zahlen $x_1, \ldots, x_n$. Durchmische die Kugeln in einem grünen Beutel und ziehe eine Kugel. Sei X die Zahl auf der gezogenen Kugel.) Zeigen Sie:

(a) Die Verteilung von X erfüllt

$$\Pr(X = a) = \frac{1}{n} \#\{i \colon x_i = a\} \qquad \text{für alle } a \in \{x_1, \ldots, x_n\},$$

wobei $\#\{i \colon x_i = a\}$ die Anzahl der Elemente der Menge ist. Im Zahlenbeispiel ist beispielsweise $\Pr(X = -1) = \frac{2}{5}$. (Diese Verteilung ist die sogenannte *empirische Verteilung* von $x_1, \ldots, x_n$.)

(b) Der Erwartungswert von X ist

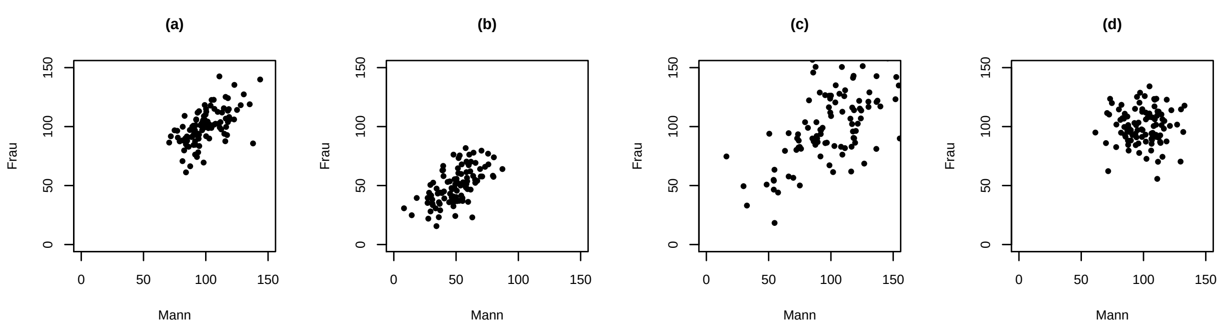
$$\mathbb{E}X = \frac{1}{n} \sum_{i=1}^{n} x_i =: \overline{x}$$

(c) Die Varianz von X ist

$$\operatorname{Var}(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

(d) Geben Sie $\mathbb{E}X$ und $\operatorname{Var}(X)$ für das Zahlenbeispiel $n = 6, x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5, x_6 = 6$ an ($\rightsquigarrow$ Würfel).

4. Aufgabe (Qualitative Aussagen aufgrund der Korrelation) 100 (simulierte) Ehepaare nahmen an einem Intelligenztest teil. Der Mittelwert des Resultats war bei beiden Geschlechtern $\overline{x}_F = \overline{x}_M = 100$ mit Standardabweichung $s_F = s_M = 15$. Wenn man jedes Ehepaar als zweidimensionale Beobachtung (x_F, x_M) auffasst, ergab sich ein Korrelationskoeffizient $\rho = 0{,}6$.
Welcher der folgenden Scatterplots passt zu diesen Beobachtungen? Begründen Sie für Bilder (a)–(d), warum sie in Frage kommen oder nicht.



Gegeben seien gepaarte Beobachtungen (x, y). Stimmen Sie den folgenden Aussagen a)–d) zu? Begründen Sie jeweils Ihre Antwort.

(a) Wenn der Korrelationskoeffizient gleich $-0{,}75$ ist, gehören unterdurchschnittliche x-Werte eher zu überdurchschnittlichen y-Werten.

(b) Wenn die y-Werte stets kleiner als die x-Werte sind, ist der Korrelationskoeffizient negativ.

(c) Wenn der Korrelationskoeffizient 1 oder -1 ist, kann man mittels des x-Werts den y-Wert exakt vorhersagen.

(d) Wenn der Korrelationskoeffizient gleich 0 ist, dann sind die den Beobachtungen zu Grunde liegenden Verteilungen unabhängig voneinander.

5. Aufgabe Jemand behauptet Ihnen gegenüber, dass es reiner Zufall ist, ob man (bei der Geburt) ein Mädchen oder einen Jungen bekommt. Entkräften Sie diese Aussage statistisch anhand folgender Daten: Im Klinikum Großhadern kamen letztes Jahr 2148 Babys zur Welt (fiktive Daten). Davon waren 1124 Jungs und 1024 Mädchen. Achten Sie auch auf eine korrekte Formulierung des Antwortsatzes.
Hinweis: Verwenden Sie den z-Test aus der Vorlesung.

6. Aufgabe Der Münzwurf ist das Parade-Beispiel einer Ja/Nein-Frage, bei der beide Antworten gleichwahrscheinlich sind.

a) Sei $X \in \{0, 1\}$ das Ergebnis eines Münzwurfes, wobei $1 :=$ Kopf und $0 :=$ Zahl. Ermitteln Sie $\mathbb{E}X$ und $\text{Var}X$.

b) Seien nun $X_1, X_2 \in \{0, 1\}$ die Ergebnisse zweier Münzwürfe, wobei $1 :=$ Kopf und $0 :=$ Zahl. Wieviele Köpfe erwarten Sie zu sehen, und wie groß ist die Varianz der Anzahl der Köpfe?

c) Stellen Sie sich vor, Sie werfen n mal eine Münze, $n \in \mathbb{N}$. Wieviele Köpfe erwarten Sie zu sehen, wie groß ist die Varianz der Anzahl der Köpfe und wie groß ist die Standardabweichung der Anzahl der Köpfe?
Hinweis: Der Abschnitt über die Binomialverteilung in der Vorlesung könnte hilfreich sein.

d) Seien $X_1, X_2, \ldots, X_n \in \{0, 1\}$ die Ergebnisse der n Münzwürfe, wobei $n \in \mathbb{N}$. Die Anzahl der Köpfe sei mit S_n bezeichnet. Was ist der Erwartungswert des empirischen Mittels $\frac{S_n}{n}$? Was ist die Standardabweichung des empirischen Mittels $\frac{S_n}{n}$? Wie verhält sich diese Standardabweichung, wenn n immer größer wird?

e) Verwenden Sie die Normalapproximation aus der Vorlesung und geben Sie an, welche Normalverteilung die Zufallsvariable S_n in etwa hat. Ersetzen Sie ??? in folgendem R code, um diese Verteilungen im Fall $n = 100$ zu plotten:

```
> plot(0:100,dbinom(0:100,size=100,p=???))
> myfun <- function(x) dnorm(x,mean=???,sd=???)
> plot(myfun,from=0,to=100,add=TRUE)
```

f) Verwenden Sie die Normalapproximation aus der Vorlesung und geben Sie an, welche Normalverteilung die Zufallsvariable $\frac{S_n}{n} - 0.5$ in etwa hat.