Datenanalyse für Naturwissenschaftler*Innen
Statistische Methoden in Theorie und Praxis **Uni Freiburg**
Fernandez Lebron, Andres (RZ: af231) WiSe 2022/23

Blatt 04 17. November 2022

Aufgabe 1 (Binomialverteilung in der Praxis)

Nehmen Sie an, dass Sie ein neues Programm schreiben und dabei in 99,2% Ihrer erstellten Programmzeilen keinen Fehler machen.

(a) Wie viele Programmzeilen müssen Sie schreiben, um mit einer Wahrscheinlichkeit von mindestens 50% mindestens einen Fehler zu machen? (4P)

(b) Wie viele Programmzeilen müssen Sie schreiben, um mit einer Wahrscheinlichkeit von mindestens 50% mindestens zwei Fehler zu machen? (4P)
Hinweis: Testen Sie mögliche Werte.

8/8

Lösung

(a) Wahrscheinlichkeit keine Fehler zu machen: $1 - p = 0.992$
Wahrscheinlichkeit eine Fehler zu machen: $p = 0.008$
$n =$Programmzeilen
$x =$Fehler
Binomialverteilung: $P(x) = \frac{n!}{(n-x)!x!}p^x(1-p)^{n-x}$

$$\begin{aligned}
P(x \geq 1) &= 1 - P(x < 1) = 1 - P(x = 0) \geq 0.5 \\
P(x = 0) &= \frac{n!}{(n-0)!0!}p^0(1-p)^{n-0} = (1-p)^n = 0.992^n \\
\implies 1 - 0.992^n &\geq 0.5 \\
0.992^n &\leq 0.5 \\
n &\geq log_{0.992}(0.5) \\
n &\geq 86.3 \text{ Programmzeilen(Also} \geq 87)
\end{aligned}$$

(b)

$$\begin{aligned}
P(x \geq 2) &= 1 - P(x < 2) = 1 - (P(x = 1) + P(x = 0)) \geq 0.5 \\
P(x = 1) &= \frac{n!}{(n-1)!1!}p^1(1-p)^{n-1} = np(1-p)^{n-1} = n0.008 \cdot 0.992^{n-1} \\
1 &- (n0.008 \cdot 0.992^{n-1} + 0.992^n) \geq 0.5 \\
\text{Für } n &= 210 : P(x \geq 2) \approx 0.5014 \\
\text{Für } n &= 209 : P(x \geq 2) \approx 0.4989 \\
\implies n &\geq 210 \text{ Programmzeilen}
\end{aligned}$$

Aufgabe 2 (Gedächtnislosigkeit der Exponentialverteilung)

Eine Wahrscheinlichkeitsverteilung P heißt gedächtnislos, wenn für alle $s, t \geq 0$ gilt:

$$P(X > s + t | X > s) = P(X > t)\,.$$

$P(X \leq x) = F(x)$ ist die kumulative Verteilungsfunktion der Wahrscheinlichkeitsverteilung.

Am Beispiel der Lebensdauerverteilung eines instabilen Teilchens, die einer Exponentialverteilung

$$f(x;\tau) = \begin{cases} \frac{1}{\tau} e^{-x/\tau} & x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

folgt, kann diese Verteilung $P(X \leq x)$ als Wahrscheinlichkeit für den Zerfall des Teilchens, bzw. $P(X > x)$ als Wahrscheinlichkeit für das Überleben des Teilchens zum Zeitpunkt x interpretiert werden. Die Gedächtnislosigkeit besagt nun, dass diese Überlebenswahrscheinlichkeit nach einer Zeit $t + s$ bezogen auf einen bestimmten Zeitpunkt s unabhängig von s selbst ist.

Beweisen Sie dies für die Exponentialverteilung.

Hinweis: Benutzen Sie die Definition der kumulativen Verteilung und stellen Sie einen Zusammenhang zwischen $P(X > x)$ und $P(X \leq x)$ her. Benutzen Sie entweder den Satz von Bayes oder überlegen Sie, zu was $P(X > s + t) \cap P(X > s)$ äquivalent ist.

Lösung

+ Rechnung für $F(x;\tau)$

Kumulative Verteilungsfunktion von der Exponentialverteilung: $F(x;\tau) = \begin{cases} 1 - e^{-x/\tau} & x \leq 0 \\ 0 & \text{sonst} \end{cases}$

$$P(X > s + t | X > s) = \frac{P(X > s + t) \cap P(X > s)}{P(X > s)}$$

Wenn $X > s + t$, dann ist bereits $X > t$ wahr:

$$\begin{aligned} P(X > s + t | X > s) &= \frac{P(X > s + t)}{P(X > s)} \\ &= \frac{1 - F(s + t;\tau)}{1 - F(s;\tau)} \quad \text{wobei } F \text{ ist kumulative Verteilungsfunktion} \\ &= \frac{e^{-(s+t)/\tau}}{e^{-s/\tau}} \\ &= e^{-t/\tau} \\ &= P(X > t) \end{aligned}$$

$\Longrightarrow$ Exponentialverteilung ist gedächtnislos

✓ 7/8

4,5/6

Aufgabe 3 (Farbige Kugeln)

Eine Kiste enthält 4 rote, 3 weiße und 2 blaue Kugeln. Eine Kugel wird zufällig herausgenommen, die Farbe notiert, und die Kugel wird wieder in die Kiste gelegt.

(a) Bestimmen Sie die Wahrscheinlichkeit für den Fall, dass Sie $n = 6$ Kugeln auf diese Weise herausnehmen und davon 3 Kugeln rot, 2 weiß und 1 blau sind. (2P)

(b) Erklären Sie die Bedeutung der einzelnen Terme, die in Ihre Rechnung eingehen. (2P)

(c) Was ist der Erwartungswert und die Varianz für die Verteilung der roten Kugeln wenn $n = 6$? Um welche Art Verteilung handelt es sich hierbei? (2P)

Lösung

(a)

$$P(\text{rot}) = \frac{4}{9}, \quad P(\text{weiss}) = \frac{3}{9}, \quad P(\text{blau}) = \frac{2}{9}$$

Die Ereignisse sind unabhängig, da die Kugeln zurückgelegt werden, also:

$$P_1 = (\frac{4}{9})^3 \cdot (\frac{3}{9})^2 \cdot (\frac{2}{9})^1 \approx 0.002$$

2/2

Es gibt $\frac{6!}{3!2!1!} = 60$ Möglichkeiten, dann ist die Wahrscheinlichkeit:

$$P = 60 \cdot P_1 \cong 60 \cdot 0.002 \cong 0.12$$

eigentlich $\approx 0,13$ ✓

(b) $P(\text{rot/weiss/blau})$ ist die Wahrscheinlichkeit, eine rote/blaue/weiße Kugel herauszunehmen.

$(\frac{4}{9})^3, (\frac{3}{9})^2, (\frac{2}{9})^1$ sind Wahrscheinlichkeite, 3 rote, 2 blaue bzw 1 weiße Kugel herauszunehmen.

P_1 ist die Wahrscheinlichkeit 3 rote, 2 blaue und 1 weiße Kugel herauszunehmen

$\frac{6!}{3!2!1!} = 60$ ist die Anzahl der Möglichkeiten, wie diese Kombination von Kugeln herausgenommen werden kann.

2/2

$P = 60 \cdot P_1$ ist die erforderliche Gesamtwahrscheinlichkeit. ✓

(c)

$$\begin{aligned}
E[\text{rot}] &= \sum f(x_i)x_i = (\frac{4}{9})^0 \cdot 0 + (\frac{4}{9})^1 \cdot 1 + (\frac{4}{9})^2 \cdot 2 + (\frac{4}{9})^3 \cdot 3 + (\frac{4}{9})^4 \cdot 4 \approx 1.26 \quad f \\
\sigma^2 &= Var(\text{rot}) = \sum (x_i - [\text{rot}])^2 f(x_i) \\
&= (0-1.26)^2(\frac{4}{9})^0 + (1-1.26)^2(\frac{4}{9})^1 + (2-1.26)^2(\frac{4}{9})^2 + (3-1.26)^2(\frac{4}{9})^3 + (4-1.26)^2(\frac{4}{9})^4 \approx 2.28 \quad f
\end{aligned}$$

0,5/2

Es handelt sich um Binomialverteilung. ✓

$\mu_r = n \cdot p_r$

$V_r = n \cdot p_r (1 - p_r)$

8/8

Aufgabe 4 (Gleichverteilte Zufallsvariablen)

(a) Zeigen Sie, dass der Mittelwert der Gleichverteilung einer Zufallsvariablen x im Intervall $[a, b]$ den Mittelwert $\mu = \frac{1}{2}(a + b)$ hat. (2P)

Nehmen Sie nun an, eine Zufallsvariable x sei gleichverteilt im Intervall $[a, b]$.

(b) Finden Sie den Erwartungswert der Zufallsvariablen $y = 1/x$. (4P)

(c) Vergleichen Sie das Ergebnis mit $1/E[x]$ für $a = 1, b = 2$. (2P)

Lösung

(a) $f(x) = \begin{cases} \frac{1}{b-a} & a \geq x \geq b \\ 0 & \text{sonst} \end{cases}$

$$\mu = \int_a^b \frac{1}{b-a} x dx = \frac{1}{b-a} [\frac{1}{2} x^2]_a^b = \frac{1}{2(b-a)} (b^2 - a^2) = \frac{1}{2} \frac{(b-a)(a+b)}{(b-a)} = \frac{1}{2}(a+b)$$

(b)

$$E[y] = E[1/x] = \int_a^b \frac{1}{x} \frac{1}{b-a} dx = \frac{1}{b-a} [ln(x)]_a^b = \frac{1}{b-a} ln(\frac{b}{a})$$

(c)

$$\begin{aligned} E[y] &= E[1/x] = \frac{1}{2-1} ln(2/1) = ln(2) \approx 0.69 \\ \frac{1}{E[x]} &= \frac{1}{\mu} = \frac{1}{\frac{1}{2}(2+1)} = \frac{2}{3} \approx 0.67 \\ E[1/x] &> \frac{1}{E[x]} \end{aligned}$$