Albert-Ludwigs-Universität Freiburg | Wintersemester 2009

Statistische Methoden der Datenanalyse

Markus Schumacher

Übung II

Matthew Beckingham und Henrik Nilsen

5.11.2009

Anwesenheitsaufgaben

Aufgabe 1 *Wechseln oder nicht?*

In einer Spielshow gibt es drei Türen. Hinter einer Tür wartet ein Gewinn, hinter den beiden anderen Türen befinden sich Nieten. Nachdem der Kandidat eine Tür ausgewählt hat, öffnet der Moderator eine Tür mit einer Niete (aber nicht die Tür, die der Kandidat gewählt hat). Der Kandidat darf erneut eine der beiden übrigen Türen wählen. Sollte er bei der ursprünglichen Tür bleiben oder besser die andere Tür wählen, um seine Gewinnchance zu erhöhen?

Aufgabe 2 *Quotient zweier Gaussverteilter Variablen*

Die Zufallsvariablen x und y seien unabhängig und identisch normalverteilt mit der folgenden Wahrscheinlichkeitsdichtefunktion:

$$f(x,y) = f(x)f(y) = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right) \tag{1}$$

Bestimmen Sie die Wahrscheinlichkeitsdichtefunktion $g(u)$ des Quotienten $u = y/x$. Tipp: Führen Sie eine geeignete Hilfsfunktion, $v(x,y) = x$ ein, transformieren $f(x,y)$ in die kombinierte Wahrscheinlichkeitsdichtefunktion $h(u,v)$ und marginalisieren Sie über v.

Aufgabe 3 *Geometrische Verteilung*

Nehmen Sie ein Zufallsexperiment mit nur zwei möglichen Ergebnissen an, z.B. beim Würfeln eine sechs zu würfeln, oder eben nicht. Die Erfolgswahrscheinlichkeit sei p. Betrachtet werden soll nun das Ereignis, dass im r-ten Versuch, also z.B. beim r-ten mal Würfeln erstmals ein Erfolg eintritt.

(i) Zeigen Sie, dass die Wahrscheinlichkeit für dieses Ereignis gegeben ist durch

$$P(r;p) = p\ (1-p)^{r-1} \tag{2}$$

(ii) Zeigen Sie, dass sich charakteristische Funktion zu dieser Zufallsverteilung ergibt zu

$$\phi(t) = \frac{p\ \exp(it)}{1-(1-p)\ \exp(it)} \tag{3}$$

(iii) Zeigen Sie mittels der charakteristischen Funktion, dass Mittelwert und Varianz dieser sogenannten geometrischen Verteilung gegeben sind durch

$$\mu = \frac{1}{p} \tag{4}$$

$$V = \frac{1-p}{p^2} \tag{5}$$

Hausaufgaben

Aufgabe 4 *Varianz und Kovarianz* **3 Punkte**

(i) Die Varianz einer Stichprobe x vom Umfang N, mit den Elementen $\{x_1,x_2,\ldots,x_N\}$, ist gegeben durch

$$V(x) = \frac{1}{N}\sum_{i=0}^{N}(x_i - \overline{x})^2 \tag{6}$$

Zeigen Sie, dass es zu

$$V(x) = \overline{x^2} - \overline{x}^2 \tag{7}$$

äquivalent ist.

(ii) Die Kovarianz der N Paare von Meßobservablen (x_1,y_1), (x_2,y_2), ... (x_N,y_N) ist gegeben durch

$$cov(x,y) = \frac{1}{N}\sum_{i=0}^{N}(x_i - \overline{x})(y_i - \overline{y}) \tag{8}$$

a) Zeigen Sie, dass es zu

$$cov(x,y) = \overline{xy} - \overline{x}\ \overline{y} \tag{9}$$

äquivalent ist.

b) Zeigen Sie, dass unter der Variabletransformation $x \to x + C$, wobei C eine Konstante ist, die Kovarianz $cov(x,y)$ konstant ist.

Aufgabe 5 *Dopingtest* **5 Punkte**

Gegeben sei ein Dopingtest für Sportler. Der Dopingtest hat eine Sensitivität (Wahrscheinlichkeit für ein positives Resultat, gegeben dass der Sportler gedopt hat) von 99% und eine Spezifität (Wahrscheinlichkeit für ein positives Resultat, gegeben dass der Sportler nicht gedopt hat) von 1%.

(i) Nehmen Sie an, dass 1% alle Sportler in einem Wettkampf leistungssteigernde Medikamente benutzen. Berechnen Sie, mit Hilfe des Bayes-Theorem, die Wahrscheinlichkeit, dass ein Sportler gedopt hat, gegeben ein positives Resultat.

(ii) Berechnen Sie nun die Wahrscheinlichkeit, dass ein Sportler gedopt hat gegeben ein positives Resultat, falls die Sensitivität des Testes 90% beträgt.

(iii) Nemen Sie nun für die Sensitivität des Tests 99% an und dass 10% alle Sportler in einem Wettkampf leistungssteigerndes Medikamente benutzen. Berechnen Sie die neue Wahrscheinlichkeit, dass ein Sportler gedopt hat, gegeben ein positives Resultat.

(iv) Nehmen Sie an, dass neben der A-Probe ein weitere Test, die B-Probe, gemacht wird. Wie gross ist die Wahrscheinlichkeit, dass der Sportler gedopt hat, wenn beide Tests positiv Ergebnisse liefern? Verwenden Sie die Werte für Sensitivität und Spezifität aus der Einleitung und (i).

Aufgabe 6 *Negative Binomialverteilung* **7 Punkte**

Betrachtet werden soll wieder ein Zufallsexperiment mit nur zwei möglichen Ergebnissen und Erfolgswahrscheinlichkeit p. Man führt es nun so lange durch, bis man insgesamt r Erfolge erzielt hat.

(i) Zeigen Sie, dass die Wahrscheinlichkeit, dabei insgesamt k Misserfolge zu haben gegeben ist durch

$$P(k;r,p) = \begin{pmatrix} k+r-1 \\ r-1 \end{pmatrix} p^r(1-p)^k \tag{10}$$

(ii) Zeigen Sie, dass sich die zugehörige charakteristische Funktion zu

$$\phi(t) = \left(\frac{p}{1-(1-p)\exp(it)}\right)^r \tag{11}$$

ergibt. Tipp: Die folgende Reihenentwicklung könnte sich als nützlich erweisen:

$$(1-x)^{-m} = 1 + mx + \frac{m(m+1)}{2!}x^2 + \frac{m(m+1)(m+2)}{3!}x^3 + \dots \qquad (12)$$

(iii) Zeigen mittels der charakteristischen Funktion, dass Mittelwert und Varianz dieser sogenannten negativen Binomialverteilung gegeben sind durch

$$\mu = r\,\frac{1-p}{p} \qquad (13)$$

$$V = r\,\frac{1-p}{p^2} \qquad (14)$$

Aufgabe 7 *Kombination zweier Gleichverteilungen* **5 Punkte**

Gegeben seien zwei unabhängige Zufallsvariablen x und y, die beide gleichverteilt sind zwischen 0 und 1, d.h. ihre Wahrscheinlichkeitsdichtefunktion ist gegeben durch

$$g(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{sonst} \end{cases} \qquad (15)$$

und analog durch $h(y)$.

(i) Benutzen Sie die Mellin-Faltung von g und h, um zu zeigen, dass die Wahrscheinlichkeitsdichtefunktion $f(z)$ von $z = xy$ gegeben ist durch:

$$f(z) = \begin{cases} -\ln z & 0 < z < 1 \\ 0 & \text{sonst} \end{cases} \qquad (16)$$

(ii) Zeigen Sie dasselbe, indem Sie eine zweite Funktion $u = x$ einführen und die gemeinsame Wahrscheinlichkeitsdichtefunktion von z und u bilden, und schliesslich über u integrieren. Tipp: Jacobi-Determinante!

(iii) Zeigen Sie, dass die Kumulativverteilung gegeben ist durch

$$F(z) = z(1-\ln z) \qquad 0 < z < 1 \qquad (17)$$