Prof. Dr. Jan Beyersmann
Matthias Weber

SS 2021
Übung: 25. Juni 2021, Abgabe: 24. Juni

Angewandte Stochastik 1 – Übungsblatt 5

Dieses Übungsblatt wird in der Übung am 25. Juni um 10:15 Uhr diskutiert. Bitte gebt eure Lösungen bis zum Donnerstag, 24. Juni um 12:00 Uhr über Moodle ab. Gebt eure Lösungen als exakt eine PDF-Datei ab, andere Dateiformate oder mehrere Dateien sind nicht erlaubt. Gebt die Lösungen nach Möglichkeit bitte in Gruppen von zwei oder drei Personen ab. Bitte beachtet auch die Hinweise zur Übung in Moodle.

Musterlösungen

Aufgabe 1 (4 Punkte)

Seien $X_1 \sim N(\mu_1, \sigma_1^2)$ und $X_2 \sim N(\mu_2, \sigma_2^2)$ unabhängige normalverteilte Zufallsvariablen wie in Folgerung 3.6.1. Zeige, dass

$$X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right).$$

Lösung:

Da X_1 und X_2 unabhängig sind, gilt die Faltungsformel für $Y = X_1 + X_2$:

$$f_Y(x) = \int_{\mathbb{R}} f_{X_1}(y) f_{X_2}(x-y)\, dy$$

Also hier:

$$\begin{aligned}
f_{X_1+X_2}(z) &= \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(z-t)\, dt \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2}\left(\frac{t-\mu_1}{\sigma_1}\right)^2\right) \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{z-t-\mu_2}{\sigma_2}\right)^2\right) dt \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2}\left(\frac{t}{\sigma_1}\right)^2\right) \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)-t}{\sigma_2}\right)^2\right) dt \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2}\left(\frac{t}{\sigma_1}\right)^2 - \frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)-t}{\sigma_2}\right)^2\right) dt \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2}\frac{(z-(\mu_1+\mu_2))^2}{\sigma_1^2+\sigma_2^2}\right) g(z),
\end{aligned}$$

mit

$$\begin{aligned}
g(z) &= \int_{-\infty}^{\infty} \exp\left(-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(t - \frac{\sigma_1^2(z-(\mu_1+\mu_2))}{\sigma_1^2+\sigma_2^2}\right)^2\right) dt \\
&= \int_{-\infty}^{\infty} \exp\left(-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\, t^2\right) dt \\
&= \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt = \frac{\sqrt{2\pi}\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}.
\end{aligned}$$

Also gilt

$$f_{X_1+X_2}(z) = \frac{1}{\sqrt{2\pi}\,\sqrt{\sigma_1^2+\sigma_2^2}} \exp\left(-\frac{1}{2}\left(\frac{z-(\mu_1+\mu_2)}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2\right).$$

Aufgabe 2 (2 + 2 + 1 = 5 Punkte)

Der Inhalt einer bestimmten Sorte Farbdosen sei auf dem Etikett mit 1000g Farbe angegeben. Die Abfüllmaschine kann diese Menge jedoch nicht exakt abfüllen, so dass die tatsächlich abgefüllte Menge (in Gramm) normalverteilt ist mit den Parametern $\mu = 1000$ und $\sigma^2 = 100$. Die verwendeten Dosen können höchstens 1020g Farbe fassen.

(a) Berechne die Wahrscheinlichkeit, dass eine Dose beim Abfüllen überläuft. Verwende dazu zunächst eine geeignete Skalierung und Normierung, um eine standardnormalverteilte Zufallsvariable zu bekommen. Verwende außerdem die unten angegebenen Werte der Verteilungsfunktion einer Standardnormalverteilung.

Lösung:
Sei $M \sim N(1000, 100)$ die abgefüllte Menge. Dann ist $X = (M - 1000)/10 \sim N(0, 1)$ (siehe Bsp. 3.6.1/1.). Also

$$P(M > 1020) = P(X > 2) = 1 - 0.977 = 0.023$$

(b) Zeige, dass für die Verteilungsfunktion $\Phi\colon \mathbb{R} \to [0, 1]$ der Standardnormalverteilung gilt: $\Phi(-x) = 1 - \Phi(x)$. Nutze dazu aus, dass die Dichte $\varphi\colon \mathbb{R} \to [0, \infty)$ der Standardnormalverteilung achsensymmetrisch ist, d.h. es gilt $\varphi(x) = \varphi(-x)$ für alle $x \in \mathbb{R}$.

Lösung:
Mit der Substitution $u = -t$ folgt:

$$\Phi(-x) = \int_{-\infty}^{-x} \varphi(t)\, dt = \int_{-\infty}^{-x} \varphi(-t)\, dt = \int_{x}^{\infty} \varphi(u)\, du = 1 - \Phi(x)$$

(c) Berechen mit Hilfe von Teil b) die Wahrscheinlichkeit, dass eine Farbdose weniger als 990g Farbe enthält. Verwende die unten angegebenen Werte der Verteilungsfunktion einer Standardnormalverteilung.

Lösung:

$$P(M < 990) = P(X < -1) = P(X > 1) = 1 - 0.841 = 0.159$$

Hinweis: Für eine standardnormalverteilte Zufallsvariable $X \sim N(0, 1)$ gilt

x	0.1	0.2	0.5	1	2	5
$F_X(x)$	0.540	0.579	0.691	0.841	0.977	1.000

Aufgabe 3 (5 Punkte)

Sei $X = (X_1, X_2)$ ein absolutstetiger Zufallsvektor mit der Dichte

$$f_X(x_1, x_2) = \begin{cases} \frac{2}{3}x_1 + \frac{4}{3}x_1x_2 + \frac{2}{3}x_2, & \text{für } (x_1, x_2) \in [0,1]^2, \\ 0, & \text{sonst.} \end{cases}$$

Prüfe, ob die Zufallsvariablen X_1 und X_2 unabhängig sind.

Lösung:

Wir berechnen die Randdichten $f_{X_1}(x)$ und $f_{X_2}(x)$ für $x \in (0,1)$:

$$f_{X_1}(x_1) = \int_0^1 f_X(x_1, t)\,dt = \int_0^1 \frac{2}{3}x_1 + \frac{4}{3}x_1 t + \frac{2}{3}t\,dt = \frac{2}{3}x_1 t + \frac{2}{3}x_1 t^2 + \frac{1}{3}t^2\Big|_0^1 = \frac{4}{3}x_1 + \frac{1}{3}$$

und offensichtlich auch:

$$f_{X_2}(x_2) = \frac{4}{3}x_2 + \frac{1}{3}$$

Es ist

$$f_{X_1}(x_1)f_{X_2}(x_2) = \left(\frac{4}{3}x_1 + \frac{1}{3}\right)\left(\frac{4}{3}x_2 + \frac{1}{3}\right) = \frac{4}{9}x_1 + \frac{16}{9}x_1x_2 + \frac{4}{9}x_2 + \frac{1}{9} \neq f_X(x_1, x_2)\,,$$

also sind die Zufallsvariablen nicht unabhängig.

Aufgabe 4 (3 Bonuspunkte)

Ein Pseudozufallszahlengenerator (englisch: pseudorandom number generator, PRNG) wird genutzt um am Computer Zahlen zu erzeugen, die als Realisierungen von auf $[0, 1]$ gleichverteilten und unabhängigen Zufallszahlen angesehen werden können. Eine einfache Klasse von Algorithmen zur Erzeugung von (Pseudo-)Zufallszahlen sind sogenannte *lineare Kongruenzgeneratoren*, die für vier Parameter $y_1, b, a \in \{0, \dots, m-1\}$ und m (genannt *Startwert, Inkrement, Faktor* und *Modul*) mit folgender rekursiven Vorschrift

$$y_i = (ay_{i-1} + b) \mod m$$

Zahlen y_i generieren. Hier bezeichnet $\mod m$ den Rest nach Division durch m.

Wir wollen nun die Güte dieser Klasse von Zufallszahlengeneratoren untersuchen. Insbesondere interessieren wir uns für die Unabhängigkeit der Realisierungen $y_1, y_2, \dots$.

Implementiere in einer Programmiersprache deiner Wahl einen linearen Kongruenzgenerator mit Parametern $a = 24298$, $b = 99991$, $m = 199017$ und $y_1 = 7324$. Generiere damit nun 180 Pseudozufallszahlen $y_1, y_2, \dots, y_{180}$ und berechne daraus zwei Listen $X = (x_1, \dots, x_{60}) = (y_1, y_4, y_7, \dots, y_{178})$ und $Z = (z_1, \dots, z_{60}) = (y_2+y_3, y_5+y_6, y_8+y_9, \dots, y_{179}+y_{180})$. Plotte nun die Punktwolke $(x_1, z_1), \dots (x_{60}, z_{60})$. Was fällt dabei auf? Würden wir dieses Verhalten erwarten, wenn die Werte y_i wirklich Realisierungen von unabhängigen Zufallsvariablen wären?

Lösung:

In R implementiert:

```
rng = function(a, b, m, y_0, n) {
        ys = c(y_0)
        for (i in 2:n) {
                ys[i] = (a * ys[i-1] + b) %% m
        }
        return(ys)
}

vals = rng(24298, 99991, 199017, 7324, 180)

X = vals[seq(1, n, 3)]
Z = vals[seq(2, n, 3)] + vals[seq(3, n, 3)]
plot(X, Z)
```

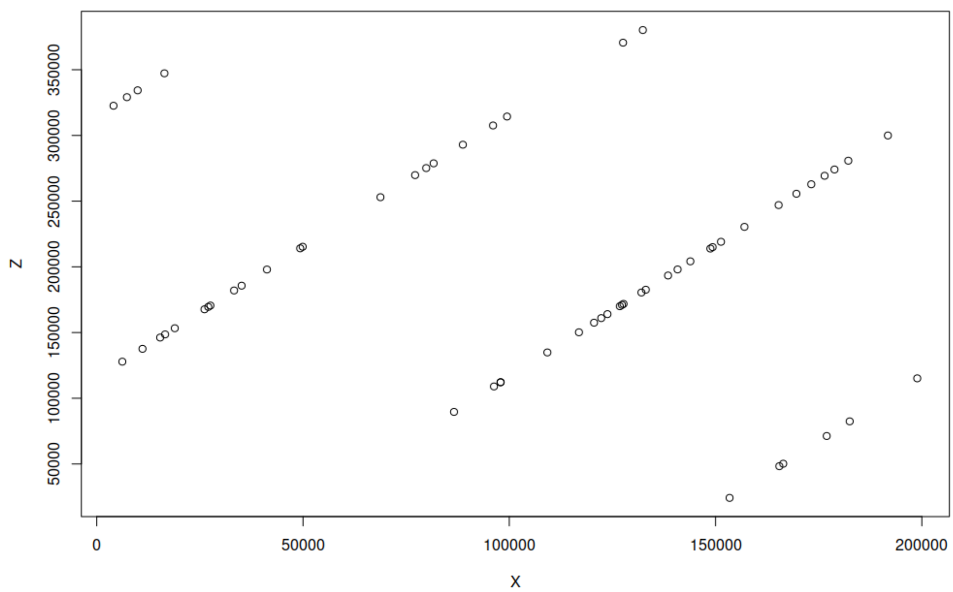

Abbildung 1: Die Punktwolke $(x_1, z_1), \dots (x_{60}, z_{60})$. Alle Punkte liegen auf parallelen Geraden, es besteht also offensichtlich eine Abhängigkeit zwischen x_i und z_i. Dann können auch die zugrundeliegenden Werte y_i nicht unabhängig sein.

Multiple Choice Aufgaben

Bei den folgenden Multiple Choice Aufgaben sind die richtigen Aussagen anzugeben. Dabei kann genau eine Aussage zutreffen, mehrere Aussagen können zutreffen oder aber keine der Aussagen trifft zu. Die Aufgabe ist richtig gelöst, wenn alle zutreffenden Aussagen angegeben wurden, jedoch keine nicht zutreffende Aussage angegeben wurde.

Für die Bewertung des Übungsblatts gibt jede richtige Antwort 1 Punkt, jede falsche Antwort gibt 1 Punkt Abzug. Je Aufgabe können minimal 0 Punkte erreicht werden.

Aufgabe 5 (6 Punkte)

Sei (X, Y) ein zweidimensionaler Zufallsvektor, der $N(\mu, \Sigma)$ verteilt ist. Gib an, welche Aussagen korrekt sind:

(a) Σ ist eine 2x2-Matrix. **Lsg:** wahr

(b) Falls X und Y unabhängig sind, dann sind die Nichtdiagonalelemente von Σ gleich null. **Lsg:** wahr

(c) Die Diagonalelemente von Σ sind kleiner Null, falls X und Y abhängig sind. **Lsg:** falsch

(d) Falls die Nichtdiagonalelemente von Σ gleich Null sind, dann ist $X + Y$ ebenfalls normalverteilt. **Lsg:** wahr

(e) Falls $X+Y$ normalverteilt ist mit Parametern m und s^2, $X+Y \sim N(m, s^2)$, so ist s^2 gleich der Summe der Nichtdiagonalelemente von Σ. **Lsg:** falsch

(f) Falls

$$\mu = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ und } \Sigma = \begin{pmatrix} 17 & 2 \\ 2 & 1 \end{pmatrix},$$

dann ist Y^2 χ^2_1-verteilt. **Lsg:** falsch