Aufgabe 1

Sei $Z = X + Y$ mit $X, Y \stackrel{iid}{\sim} \mathcal{E}(\lambda)$. Zeigen Sie, dass Z Gamma-verteilt ist mit $Z \sim \mathcal{G}(\alpha = 2, \beta = \lambda)$.

$f_X(x) = \lambda e^{-\lambda x} \cdot I_{(0,\infty)}(x)$, $f_Y(y) = \lambda e^{-\lambda y} \cdot I_{(0,\infty)}(y)$ $\qquad f_{\Gamma(\alpha,\beta)}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \cdot e^{-\beta x}$

$T_X = T_Y = T_Z = (0,\infty)$

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\, dx \cdot I_{(0,\infty)}(z) = \int_{-\infty}^{\infty} \lambda e^{-\lambda x} \cdot I_{(0,\infty)}(x) \cdot \lambda e^{-\lambda(z-x)} \cdot I_{(0,\infty)}(z-x)\, dx \cdot I_{(0,\infty)}(z)$$

$$= \int_0^z \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(z-x)} dx \cdot I_{(0,\infty)}(z) = \lambda^2 \cdot e^{-\lambda z} \cdot \int_0^z 1\, dx \cdot I_{(0,\infty)}(z) = \lambda^2 \cdot z \cdot e^{-\lambda z} \cdot I_{(0,\infty)}(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot z^{\alpha-1} \cdot e^{-\beta \cdot z} \cdot I_{(0,\infty)}(z)$$ mit $\beta = \lambda$ und $\alpha = 2$.

$\Rightarrow Z \sim \Gamma(\alpha = 2, \beta = \lambda)$

Aufgabe 2

Sei die gemeinsame Dichte von (X, Y) gegeben als

$$f(x,y) = \begin{cases} \frac{1}{x} & \text{für} \quad 0 \le y \le x \le 1 \\ 0 & \text{sonst} \end{cases}$$

(a) Was ist der Träger von (X, Y)?

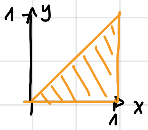
(b) Zeigen Sie dass die Randverteilung von X eine Gleichverteilung auf $[0, 1]$ ist.

(c) Bestimmen Sie die Randverteilung von Y.

(d) Zeigen Sie dass die bedingte Verteilung von $Y|X = x$ eine Gleichverteilung auf $[0, x]$ ist.

(e) Bestimmen Sie die bedingte Verteilung von $X|Y = y$.

a) $T_{X,Y} := \{(x,y) \in [0,1]^2 \mid y \le x\}$



b) $f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\, dy = \int_{\mathbb{R}} \frac{1}{x} \cdot I_{[0,x]}(y)\, dy \cdot I_{[0,1]}(x) = \frac{1}{x} \int_0^x 1\, dy \cdot I_{[0,1]}(x) = \frac{1}{x} \cdot x \cdot I_{[0,1]}(x) = 1 \cdot I_{[0,1]}(x)$

c) $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\, dy = \int_{\mathbb{R}} \frac{1}{x} \cdot I_{[y,1]}(x)\, dx \cdot I_{[0,1]}(y) = \int_y^1 \frac{1}{x}\, dx \cdot I_{[0,1]}(y) = (\ln(1) - \ln(y)) \cdot I_{[0,1]}(y) = \ln(\frac{1}{y}) \cdot I_{(0,1)}(y)$

d) Sei $x \in [0,1]$. $f_{Y|X}(x,y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\frac{1}{x} \cdot I_{[0,x]}(y) \cdot \cancel{I_{[0,1]}(x)}}{\cancel{I_{[0,1]}(x)}} = \frac{1}{x} \cdot I_{[0,x]}(y)$

e) Sei $y \in (0,1)$. $f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{x} \cdot \cancel{I_{(0,1)}(y)} \cdot I_{[y,1]}(x)}{\ln(\frac{1}{y}) \cdot \cancel{I_{(0,1)}(y)}} = \frac{1}{x \cdot \ln(\frac{1}{y})} \cdot I_{[y,1]}(x)$

Sei $y \in (0,1)$. $F_{X|Y}(X \le x \mid Y = y) = \int_{-\infty}^{x} \frac{1}{u \cdot \ln(\frac{1}{y})} \cdot I_{[y,1]}(u)\, du = \begin{cases} 0 & , \text{für } x < y \\ 1 & , \text{für } x \ge 1 \\ \frac{1}{\ln(\frac{1}{y})} (\ln(x) - \ln(y)) & , \text{für } x \in [y,1) \end{cases} = \begin{cases} 0 & , \text{für } x < y \\ 1 & , \text{für } x \ge 1 \\ 1 - \frac{\ln(x)}{\ln(y)} & , \text{für } x \in [y,1) \end{cases}$

Aufgabe 3

Die gemeinsame Wahrscheinlichkeitsfunktion von X und Y sei

$$f(x,y) = \begin{cases} \exp(-2\lambda)\frac{\lambda^{x+y}}{x!y!} & x,y \in \mathbb{N}_0^+ \\ 0 & \text{sonst} \end{cases}$$

(a) Bestimmen Sie die Randverteilungen von X und Y

(b) Bestimmen Sie die bedingten Verteilungen von $X|Y = y$ und $Y|X = x$

(c) Was folgt aus den Ergebnissen der vorigen Teilaufgaben für (X,Y)?

a) $$f_X(x) = \sum_{y=-\infty}^{\infty} f_{X,Y}(x,y) = \sum_{y=0}^{\infty} \exp(-2\lambda)\cdot\frac{\lambda^{x+y}}{x!y!} = \exp(-2\lambda)\cdot\frac{1}{x!}\cdot\lambda^x\cdot\sum_{y=0}^{\infty}\frac{\lambda^y}{y!} = \exp(-2\lambda)\cdot\frac{1}{x!}\cdot\lambda^x\cdot e^{\lambda} = \frac{\lambda^x\cdot e^{-\lambda}}{x!} \Rightarrow X \sim \mathcal{P}(\lambda)$$

Wegen Symmetrie gilt $Y \sim \mathcal{P}(\lambda)$.

b) $$f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\exp(-2\lambda)\cdot\frac{\lambda^{x+y}}{x!y!}}{\frac{\lambda^y\cdot e^{-\lambda}}{y!}} = \frac{e^{-\lambda}\cdot\lambda^x}{x!} \Rightarrow X|Y \sim \mathcal{P}(\lambda).$$ Wegen Symmetrie gilt $Y|X \sim \mathcal{P}(\lambda)$.

c) $f_{X|Y}(x,y) = f_X(x) \wedge f_{Y|X}(x,y) = f_Y(y)$. $\Rightarrow$ X und Y sind unabhängig.

Aufgabe 4

Der Zwedschgen-Alfons betreibt einen Obststand vor der Uni. Sei $N \sim \mathcal{P}(\lambda)$ die Poisson-verteilte Anzahl an Kund:innen an einem gegebenen Tag. Seien die Umsätze U_i für die i-te Person des Tages unabhängig und identisch Gamma-verteilt: $U_i \overset{iid}{\sim} \mathcal{G}(\alpha,\beta)$

(a) Helfen Sie dem Zwedschgen-Alfons den Erwartungswert seines täglichen Gesamtumsatzes zu bestimmen.

(b) Helfen Sie dem Zwedschgen-Alfons die Varianz seines täglichen Gesamtumsatzes zu bestimmen.

(c) Erfahrungsgemäß kommen im Mittel 120 Kund:innen pro Tag zum Obststand und geben jeweils im Mittel 10 Euro aus, mit einer Standardabweichung von $\sqrt{10}$ Euro. Bestimmen Sie (approximativ) die Wahrscheinlichkeit dafür, dass der Zwedschgen-Alfons an einem gegebenen Tag weniger als 1000 Euro Umsatz macht.

(d) Verifizieren Sie ihr Ergebnis aus der vorherigen Teilaufgabe mit einem geeigneten Simulations-Experiment in R.

Hinweis: Die Summe von n unabhängigen $\mathcal{G}(\alpha,\beta)$-Zufallsvariablen ist $\mathcal{G}(n\alpha,\beta)$-verteilt.

Gesamtumsatz $= U = \sum_{i=0}^{N} U_i \overset{\text{Hinweis}}{\sim} \Gamma(N\alpha,\beta)$

a) $\mathbb{E}[U|N=n] = \sum_{i=0}^{n} E(U_i) = n\cdot\frac{\alpha}{\beta}$. $\mathbb{E}[U] = \sum_{i=0}^{\infty} E(Y|N=i)\cdot P(N=i) = \sum_{i=0}^{\infty} i\cdot\frac{\alpha}{\beta}\cdot\frac{e^{-\lambda}\cdot\lambda^i}{i!} = \frac{\alpha}{\beta}\cdot e^{-\lambda}\cdot\sum_{i=1}^{\infty}\frac{\lambda^i}{(i-1)!} = \frac{\alpha}{\beta}\cdot e^{-\lambda}\cdot\lambda\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = \frac{\alpha}{\beta}\cdot\lambda$

Satz vom iterierten Erwartungswert: $\mathbb{E}[U] = \mathbb{E}_N[\mathbb{E}(U|N)] = \mathbb{E}_N\left[\frac{N\cdot\alpha}{\beta}\right] = \frac{\alpha}{\beta}\cdot\mathbb{E}_N[N] = \frac{\alpha\cdot\lambda}{\beta}$

b) Satz von der totalen Varianz: $Var(X) = E(Var(X|Z)) + Var(E(X|Z))$

$$Var(U) = \mathbb{E}(\underbrace{Var(U|N)}_{\frac{N\alpha}{\beta^2}}) + Var(\underbrace{\mathbb{E}(U|N)}_{\frac{N\alpha}{\beta}}) = \frac{\alpha}{\beta^2}\mathbb{E}(N) + \frac{\alpha^2}{\beta^2}\cdot Var(N) = \frac{\lambda\alpha(1+\alpha)}{\beta^2}$$

c) Gesamtumsatz an Tag i: $U = U_1 + U_2 + U_3 + \ldots + U_N$.

$\mathbb{E}[N] = 120 \Rightarrow \lambda = 120$, $\mathbb{E}[U_i] = 10 = \frac{\alpha}{\beta} \Rightarrow \mathbb{E}[U] = \frac{\alpha}{\beta}\cdot\lambda = 10\cdot 120 = 1200$

$Var(N) = \lambda = 120$, $Var(U_i) = 10 = \frac{\alpha}{\beta^2} \Rightarrow \alpha = 10, \beta = 1 \Rightarrow Var(U) = \frac{\lambda\alpha(1+\alpha)}{\beta^2} = \frac{120\cdot 10\cdot(1+10)}{1} = 13200$

$U \overset{a}{\sim} \mathcal{N}(1200, 13200) \Rightarrow \mathbb{P}(U \leq 1000) = \mathbb{P}\left(\tilde{U} \leq \frac{1000-1200}{\sqrt{13200}}\right) = \Phi(-1.74) = 1 - \Phi(1.74) \approx 0.041$

d)

```r
# inputs:
#   n: simulate how many days
#   lambda: daily rate of customers
#   alpha, beta: Gamma-parameters for sales per customer
# output:
#   total sales per day
simulate_sales <- function(n, lambda, alpha, beta) {
  # check inputs:
  checkmate::assert_count(n)
  checkmate::assert_number(lambda, lower = 0)
  checkmate::assert_number(alpha, lower = 0)
  checkmate::assert_number(beta, lower = 0)

  daily_customers <- rpois(n, lambda = lambda)
  # rgamma is vectorized, so we can hand over a different <shape> for each day:
  daily_sales <- rgamma(n = n, shape = daily_customers * alpha, rate = beta)
  daily_sales
}
```

Sample distribution 100000 times, with parameters from exercise:

```r
set.seed(1312)
sales <- simulate_sales(n = 1e6, lambda = 120, alpha = 10, beta = 1)
```

Simulated probability of sales below 1000:

```r
mean(sales < 1000)
```

```
## [1] 0.037436
```

Compare approximation via CLT:

```r
pnorm(1000, mean = 120 * 10, sd = sqrt(120 * 10 * 11))
```

```
## [1] 0.04086138
```

```r
# extra: compare simulated distribution (black) with asymptotic Gaussian (dashed red):
plot(density(sales), main = "Dichte des simulierten Tagesumsatzes")
curve(dnorm(x, m = 1200, sd = sqrt(13200)), col = 2, lty = 2, add = TRUE)
```