Übungsblatt 9 zur Vorlesung "Einführung in die Statistik"

Olivier Warin

17. November 2012

Aufgabe 60 [Ablesen von Wahrscheinlichkeiten aus Tabellen, Z-Transform]

a) Sei X eine $\mathcal{N}(2,9)$-Zufallsgrösse. Mit Hilfe der Z-Transform (Aufgabe 63) und der entsprechenden Tabelle (Krengel) berechnen wir:

$$\begin{aligned} P[-2<X<3] &= P\left[\frac{-2-2}{\sqrt{9}}<\frac{X-2}{\sqrt{9}}<\frac{3-2}{\sqrt{9}}\right] \\ &\overset{\text{Aufg}}{\underset{63}{=}} P[-4/3<\mathcal{N}(0,1)<1/3] \\ &= P[\mathcal{N}(0,1)<1/3]-P[\mathcal{N}(0,1)\leqslant -4/3] \\ &= P[\mathcal{N}(0,1)<1/3]-(1-P[\mathcal{N}(0,1)\leqslant 4/3]) \\ &\underset{\text{Krengel}}{\doteq} 0.6293-(1-0.9082) \\ &= 0.5375. \end{aligned}$$

Alternativ berechnen wir mit R:

$$\begin{aligned} P[-2<X<3] &= P[X<3]-P[X\leqslant -2] \underset{\text{R}}{\doteq} \texttt{pnorm(3,2,sqrt(9))-pnorm(-2,2,sqrt(9))} \\ &\underset{\text{R}}{\doteq} 0.5393474. \end{aligned}$$

b) Sei X eine $\mathcal{N}(20,25)$-Zufallsgrösse. Mit Hilfe der Z-Transform (Aufgabe 63) und der entsprechenden Tabelle (Krengel) berechnen wir:

$$P[X>19] \overset{\text{Aufg}}{\underset{63}{=}} P\left[\mathcal{N}(0,1)>\frac{19-20}{\sqrt{25}}\right] \overset{\text{Sym-}}{\underset{\text{metrie}}{=}} P[\mathcal{N}(0,1)\leqslant 0.2] \underset{\text{Krengel}}{\doteq} 0.5793.$$

Alternativ berechnen wir mit R:

$$P[X>19] = \underset{\text{R}}{\doteq} \texttt{pnorm(19,20,sqrt(25),lower.tail=FALSE)} \underset{\text{R}}{\doteq} 0.5792597.$$

c) Sei X eine χ^2_3-Zufallsgrösse. Mit Hilfe der entsprechenden Tabelle (Krengel) berechnen wir:

$$P[X\geqslant 6.25] = 1-P[X<6.25] \underset{\text{Krengel}}{\doteq} 1-0.900 = 0.1.$$

Alternativ berechnen wir mit R:

$$P[X\geqslant 6.25] \underset{\text{R}}{\doteq} \texttt{pchisq(6.25,3,lower.tail=FALSE)} \doteq 0.1000608.$$

Aufgabe 61 [Halbwertszeit, Median, Erwartungswert bei Exp]

Y habe eine Exponentialverteilung mit Parameter $\lambda>0$; d.h. die Dichte sei $f(x)=\lambda e^{-\lambda x}$ für $x\geqslant 0$ und sonst gleich 0.

Der Erwatungswert $E[Y]$ ist nach Aufgabe 49 a) gegeben durch $E[Y]=1/\lambda$.

Sei $m\in\mathbb{R}$ der Median von $E[Y]$. Es gilt also

$$1/2 = P[X\leqslant m] = \int_0^m \lambda e^{-\lambda x}dx = 1-e^{-\lambda m}.$$

Beachte dazu, dass klar gilt $m > 0$. Wir schliessen:

$$m = \frac{\log(2)}{\lambda} = E[Y]\log(2).$$

Die Halbwertszeit ist (im Modell für radioaktiven Atomzerfall) die (erwartete) Zeit, bis die Hälfte aller Atome zerfallen ist. Dies entspricht genau dem Median. Eine exakte Erklärung dafür folgt in Kapitel 5. Also ist die Halbwertszeit gegeben durch

$$\frac{\log(2)}{\lambda} = E[Y]\log(2).$$

Das Verhältnis zwischen Halbwertszeit und Erwartungswert ist daher genau $\log(2)$.

Aufgabe 62 [Gedächtnislosigkeit von Ge und Exp]

a) Sei G eine geometrisch verteilte Zufallsgrösse mit Parameter $p \in (0,1)$. Seien weiter n und m zwei natürliche Zahlen mit $n > m > 0$.

BEHAUPTUNG: Es gilt:

$$P[G > n | G > m] = P[G > n - m].$$

Beweis: Für eine natürliche Zahl $\ell > 0$ gilt:

$$P[G > \ell] = \sum_{k=\ell+1}^{\infty} P[G = k] = \sum_{k=\ell+1}^{\infty} p(1-p)^{k-1} = (1-p)^{\ell} \sum_{k=1}^{\infty} p(1-p)^{k-1} \overset{\text{Aufg}}{\underset{\text{28 a)}}{=}} (1-p)^{\ell}. \qquad (*)$$

Damit schliessen wir:

$$P[G > n | G > m] \overset{\text{Def}}{\underset{1.5}{=}} \frac{P[G > n, G > m]}{P[G > m]} \overset{n \geq m}{=} \frac{P[G > n]}{P[G > m]} \overset{(*)}{=} (1-p)^{n-m} \overset{(*)}{=} P[G > n - m].$$

■

b) Sei T eine exponential verteilte Zufallsgrösse mit Parameter $\lambda > 0$. Seien weiter t und s zwei reelle Zahlen mit $t > s > 0$.

BEHAUPTUNG: Es gilt:

$$P[T > t | T > s] = P[T > t - s].$$

Beweis: Für eine reelle Zahl $r > 0$ gilt:

$$P[T > r] = \int_r^{\infty} \lambda e^{-\lambda x} dx = e^{-\lambda r}. \qquad (**)$$

Wir schliessen:

$$P[T > t | T > s] \overset{\text{Def}}{\underset{1.5}{=}} \frac{P[T > t, T > s]}{P[T > s]} \overset{t \geq s}{=} \frac{P[T > t]}{P[T > s]} \overset{(**)}{=} \frac{e^{-\lambda t}}{e^{-\lambda s}} = e^{-\lambda(t-s)} \overset{(**)}{=} P[T > t - s].$$

■

Aufgabe 63 [Z-Transform]

Sei X eine $\mathcal{N}(\mu, \sigma^2)$-Zufallsgrösse.

BEHAUPTUNG: Die Zufallsgrösse

$$Z := \frac{X - \mu}{\sigma}$$

hat eine $\mathcal{N}(0,1)$-Verteilung.

Beweis: Seien F_X und F_Z die Verteilungsfunktionen von X bzw. Z. Nun gilt:

$$F_Z(a) = P[Z \leqslant a] = P\left[\frac{X - \mu}{\sigma} \leqslant a\right] = P[X \leqslant \sigma a + \mu] = F_X(\sigma a + \mu).$$

Seien weiter f_X und f_Z die Dichtefunktionen von X bzw. Z. Wir schliessen:

$$f_Z(a) = \frac{d}{da} F_Z(a) = \frac{d}{da} F_X(\sigma a + \mu) = f_X(\sigma a + \mu) \cdot \sigma = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\sigma a + \mu - \mu)^2}{2\sigma^2}} \cdot \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}},$$

also hat Z eine $\mathcal{N}(0,1)$-Verteilung.

$\blacksquare$

Aufgabe 64 [Bin + Bin = Bin wenn p gleich]
Seien X eine $\text{Bin}(n_1, p)$-verteilte und Y eine $\text{Bin}(n_2, p)$-verteilte Zufallsgrösse.

Behauptung: $X + Y$ hat eine $\text{Bin}(n_1 + n_2, p)$-Verteilung.

Beweis: Sei $k \in \{0, 1, \ldots, n_1 + n_2\}$.

Wir zählen zunächst die Anzahl Möglichkeitenkeiten k Elemente aus n_1+n_2 auszuwählen. Einerseits ist diese Anzahl bekanntlich genau $\binom{n_1+n_2}{k}$. Andererseits müssen wir, um k Elemente aus n_1+n_2 auszuwählen, erst i Elemente aus n_2 und dann $k-i$ Elemente aus n_1 auswählen. Wobei i natürlich alle möglichen Werte durchläuft. Dieses kombinatorische Argument liefert:

$$\binom{n_1+n_2}{k} = \sum_i \binom{n_1}{k-i}\binom{n_2}{i}. \tag{$*$}$$

Mit der Formel aus 4.4.1.1 berechnen wir nun:

$$\begin{aligned}
P[X+Y=k] &= \sum_i P[X=k-i]P[Y=i] \\
&= \sum_i \binom{n_1}{k-i} p^{k-i}(1-p)^{n_1-(k-i)} \binom{n_2}{i} p^i (1-p)^{n_2-i} \\
&= p^k (1-p)^{n_1+n_2-k} \sum_i \binom{n_1}{k-i}\binom{n_2}{i} \\
&\overset{(*)}{=} \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k}.
\end{aligned}$$

Also hat $X + Y$ eine $\text{Bin}(n_1 + n_2, p)$-Verteilung.

$\blacksquare$

Aufgabe 65 [$U[0, 1] + U[0, 1] \neq U[0, 2]$]
Seien X_1 und X_2 zwei unabhängige $U[0, 1]$-Zufallsgrössen mit Dichten f_{X_1} bzw. f_{X_2}. Sei weiter $f_{X_1+X_2}$ die Dichte von $X_1 + X_2$. Mit 4.4.1.2. folgt nun:

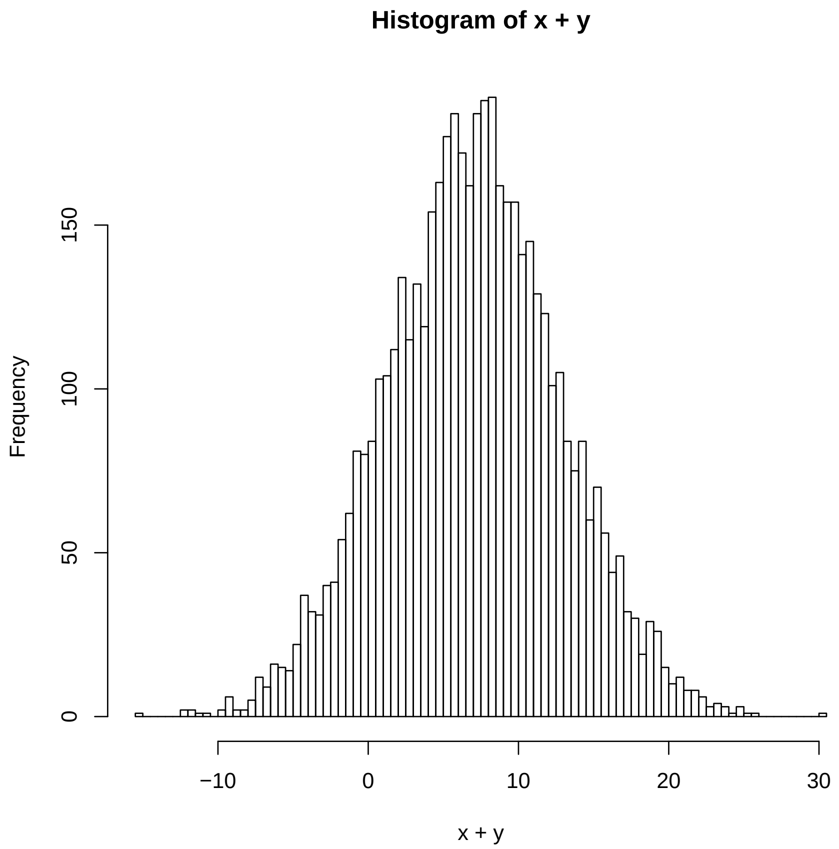
$$\begin{aligned}
f_{X_1+X_2}(a) &= \int_{-\infty}^{\infty} f_{X_1}(a-x) f_{X_2}(x)dx = \int_0^1 f_{X_1}(a-x) f_{X_2}(x)dx = \int_0^1 f_{X_1}(a-x)dx \\
&\overset{y=a-x}{=} \int_{a-1}^{a} f_{X_1}(y)dy = \begin{cases} a, & \text{falls } 0 \leqslant a \leqslant 1 \\ 2-a, & \text{falls } 1 \leqslant a \leqslant 2 \\ 0, & \text{sonst.} \end{cases}
\end{aligned}$$

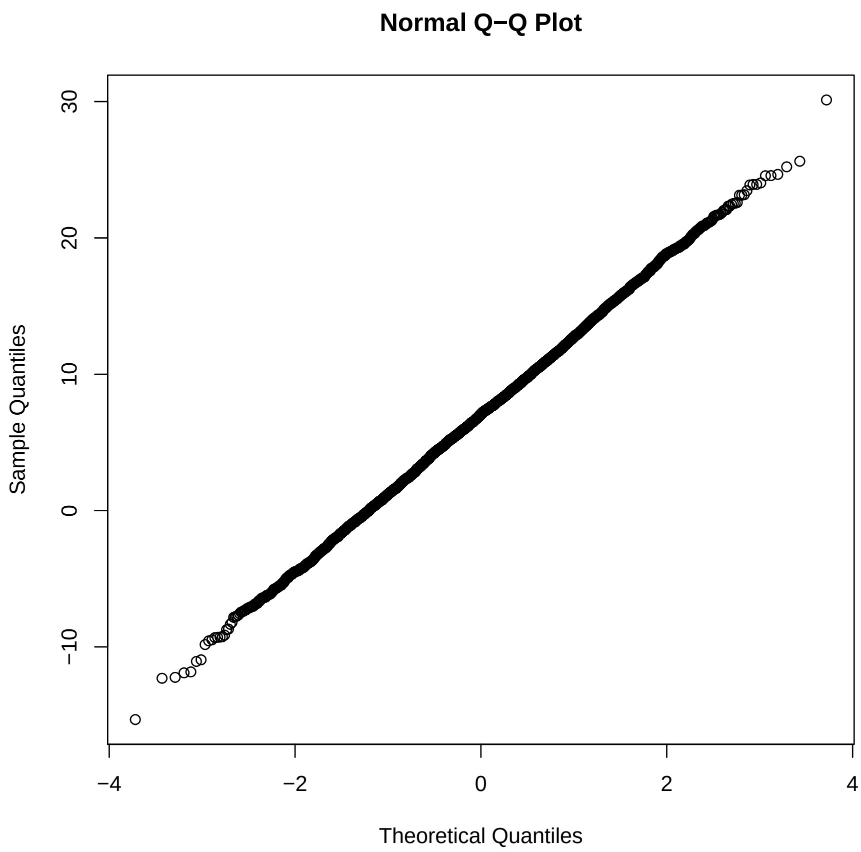
Aufgabe 66 [R, Summe von normalverteilten Zufallsgrössen]

a)

```
> x <- rnorm(5000,3,sqrt(9))
> y <- rnorm(5000,4,sqrt(25))
> mean(x+y) #Sollte etwa 3+4=7 sein
 [1] 7.038263
> var(x+y) #Sollte etwa 9+25=34 sein
 [1] 34.7063
```

Diese R-Session liefert die folgenden beiden Plots:

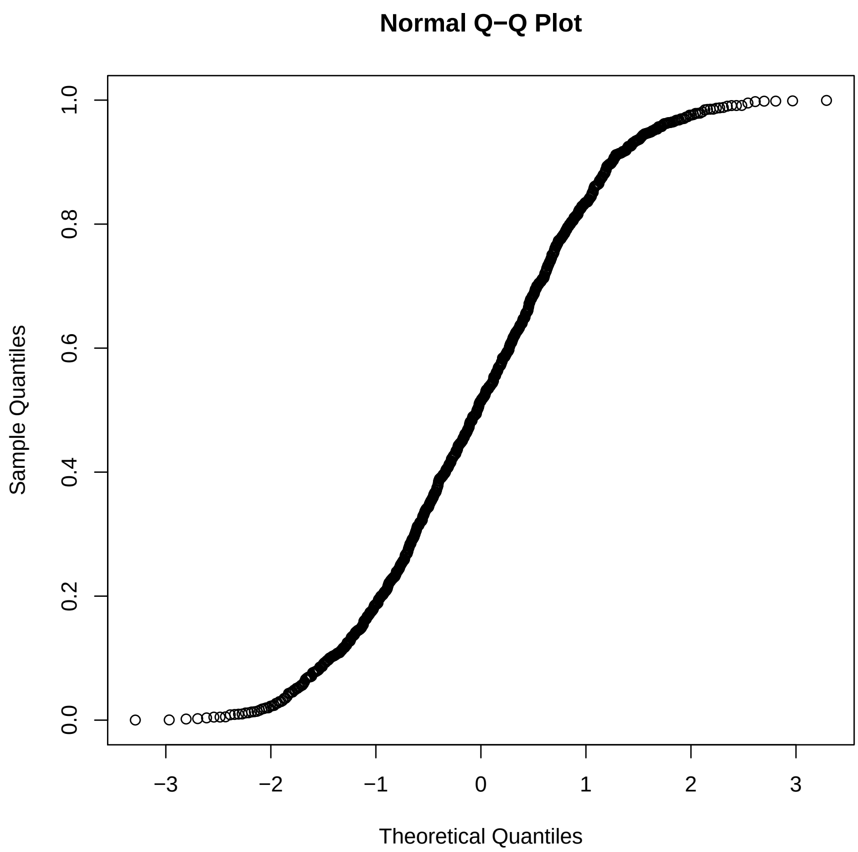

Histogram of x + y
Frequency
0
50
100
150
−10
0
10
20
30
x + y


Normal Q−Q Plot
Sample Quantiles
−10
0
10
20
30
−4
−2
0
2
4
Theoretical Quantiles

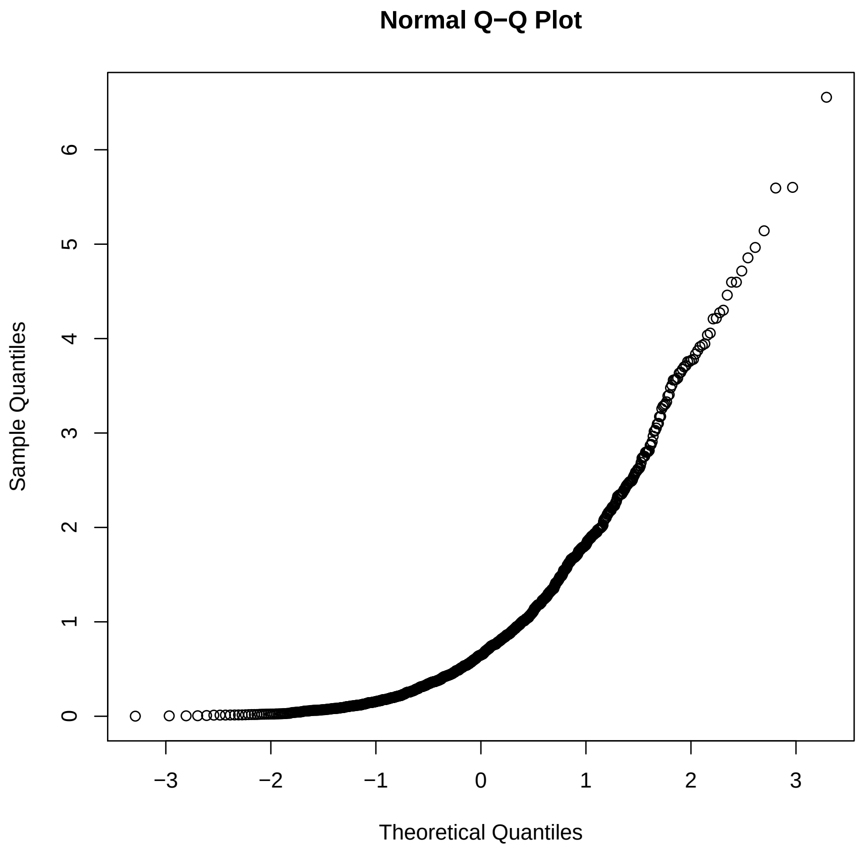
b) Der Befehl `qqnorm` plottet die Quantile der empirischen Verteilung einer Stichprobe gegen die theoretischen Quantile der $\mathcal{N}(0,1)$-Verteilung. Dies bedeutet, je näher die Punkte auf einer Gerade liegen desto eher stammt die Stichprobe aus einer normalverteilten Zufallsgrösse. Daher ist es in a) plausibel, dass `x+y` aus einer normalverteilten Zufallsgrösse stammt.

Hier noch zwei Beispiele, mit Stichproben die nicht aus einer normalverteilten Zufallsgrösse stammen:

Der Befehl `qqnorm(runif(1000))` liefert den folgenden Plot:



Der Befehl `qqnorm(rexp(1000))` liefert den folgenden Plot:



Aufgabe 67 [Standardmodell aus der Finanzmathematik]
In der Finanzmathematik wird das Modell von Samuelson (in diesem Modell gilt die berühmte Formel von Black-Scholes zur Bewertung einer europäischen Call-Option) der Kurs einer Aktie mit folgender Formel modelliert:

$$S_t := S_0 \exp\left((\mu - \sigma^2/2)t + \sigma X_t\right).$$

Dabei ist μ die Drift und σ die Volatilität. X_t ist eine sogenannte Standard-Brown'sche Bewegung. Dies impliziert insbesondere, dass X_t zur Zeit t eine $\mathcal{N}(0,t)$-Zufallsgrösse ist. Wir nehmen weiter an, dass gilt: $S_0 = 1$ (Wert der Aktie Heute), $\mu = 0.05$, $\sigma = 0.1$.

Nun gilt:

$$\begin{aligned}
P[S_1 \in [1.1, 1.4]] &= P\left[1.1 \leqslant S_0 \exp\left((\mu - \sigma^2/2) + \sigma X_1\right) \leqslant 1.4\right] \\
&\overset{S_0 = 1}{=} P\left[\log(1.1) \leqslant \mu - \sigma^2/2 + \sigma X_1 \leqslant \log(1.4)\right] \\
&\underset{\sigma = 0.1}{\overset{\mu = 0.05}{=}} P[\log(1.1) \leqslant 0.045 + 0.1 \cdot X_1 \leqslant \log(1.4)] \\
&= P[10\log(1.1) - 0.45 \leqslant X_1 \leqslant 10\log(1.4) - 0.45] \\
&= P[X_1 \leqslant 10\log(1.4) - 0.45] - P[X_1 \leqslant 10\log(1.1) - 0.45] \\
&\underset{\mathrm{R}}{\doteq} \texttt{pnorm(10*log(1.4)-0.45)-pnorm(10*log(1.1)-0.45)} \doteq 0.3056663.
\end{aligned}$$

Aufgabe 68 [Dichte der Gamma(n, λ)-Verteilung]
BEHAUPTUNG: Die Dichtefunktion f der Gamma(n, λ)-Verteilung ist gegeben durch:

$$f(y) = \frac{y^{n-1} e^{-\lambda y} \lambda^n}{\Gamma(n)}, \quad \text{falls } y \geqslant 0$$

und gleich Null sonst.

Beweis: Für $n = 1$ ist die Gamma(n, λ)-Verteilung genau eine Exp(λ)-Verteilung und es gilt daher

$$f(y) = \lambda e^{-\lambda y} = \frac{y^{n-1} e^{-\lambda y} \lambda^n}{\Gamma(n)}, \quad \text{falls } y \geqslant 0$$

und gleich Null sonst. Also stimmt die Behauptung für $n = 1$.

Nehmen wir nun an, dass $n > 1$ und die Behauptung für $n - 1$ bereits gezeigt ist. Seien X und Y zwei unabhänige Zufallsgrössen, wobei X eine Exp(λ)- und Y eine Gamma($n-1, \lambda$)-Verteilung besitze. Weiter bezeichnen wir mit f_X und f_Y die Dichtefunktionen von X bzw. Y. Es gilt also (für $y \geqslant 0$):

$$f_X(y) \; = \; \lambda e^{-\lambda y} \quad \text{und} \quad f_Y(y) \overset{\text{I.A.}}{=} \frac{y^{n-2} e^{-\lambda y} \lambda^{n-1}}{\Gamma(n-1)}$$

Ausserdem gilt $f_X(y) = f_Y(y) = 0$ für $y < 0$.

Die Zufallsgrösse $X + Y$ hat nun nach 4.3.3 eine Gamma(n, λ)-Verteilung. Also hat $X + Y$ die Dichte f. Mit 4.4.1.2 schliessen wir daher (für $y \geqslant 0$):

$$\begin{aligned} f(y) &= \int_{-\infty}^{\infty} f_X(y-x) f_Y(x) dx = \int_0^y \frac{\lambda e^{-\lambda(y-x)} x^{n-2} e^{-\lambda x} \lambda^{n-1}}{\Gamma(n-1)} dx \\ &= \frac{e^{-\lambda y} \lambda^n}{\Gamma(n-1)} \int_0^y x^{n-2} dx = \frac{e^{-\lambda y} \lambda^n y^{n-1}}{\Gamma(n-1) \cdot (n-1)} = \frac{y^{n-1} e^{-\lambda y} \lambda^n}{\Gamma(n)}. \end{aligned}$$

Ausserdem gilt für $y < 0$ klar $f(y) = 0$.

■