

Vorlesung 10a

Zweistufigkeit (Teil 2);

bedingte Verteilung,
bedingte Wahrscheinlichkeiten

1. Addieren von unabhängigen ZV'en
 - zweistufig aufgefasst:

Der diskrete Fall

(Buch S. 92)

Y und Z seien unabhängige \mathbb{Z} -wertige Zufallsvariable.

Wie ist $Y + Z$ verteilt?

Wir können $Y + Z$ auffassen als
zweite Stufe eines Zufallsexperiments.

Die erste Stufe ist Y . Gegeben $\{Y = a\}$ ist

$$X_2 := Y + Z$$

so verteilt wie $a + Z$.

$$\begin{aligned}\mathbf{P}(Y = a, Y + Z = b) &= \mathbf{P}(Y = a, a + Z = b) \\ &= \mathbf{P}(Y = a, Z = b - a) \\ &= \mathbf{P}(Y = a) \mathbf{P}(Z = b - a) = \mathbf{P}(Y = a) \mathbf{P}(a + Z = b)\end{aligned}$$

Summation über a ergibt die “totale Wahrscheinlichkeit”:

$$\mathbf{P}(Y + Z = b) = \sum_a \mathbf{P}(Y = a) \mathbf{P}(Z = b - a)$$

(Zerlegung von $\mathbf{P}(Y + Z = b)$ nach den Ausgängen von Y ,
“Zerlegung nach dem ersten Schritt”)

$$\mathbf{P}(Y + Z = b) = \sum_a \mathbf{P}(Y = a) \mathbf{P}(Z = b - a)$$

Beispiel:

Y, Z unabhängig und $\text{Geom}(p)$ -verteilt

$$\begin{aligned} \mathbf{P}(Y + Z = b) &= \sum_{a=1}^{b-1} pq^{a-1} pq^{b-a-1} \\ &= (b-1)p^2q^{b-2}, \quad b = 2, 3, \dots \end{aligned}$$

Dies sind die Gewichte der sogenannten

negativen Binomialverteilung mit Parametern $2, p$

Diese ist die Verteilung der Anzahl der Versuche in einem p -Münzwurf bis einschließlich zum zweiten Erfolg.

2. Addieren von unabhängigen ZV'en

– zweistufig aufgefasst:

Der Fall mit Dichten

(Buch S. 92)

Im diskreten Fall hatten wir

$$\begin{aligned}\mathbf{P}(Y = a, Y + Z = b) &= \mathbf{P}(Y = a, a + Z = b) \\ &= \mathbf{P}(Y = a) \mathbf{P}(a + Z = b) \\ &= \mathbf{P}(Y = a) \mathbf{P}(Z = b - a)\end{aligned}$$

Haben Y und Z die Dichten $f(y) dy$ und $g(z) dz$,
so bekommt man analog

die gemeinsame Dichte von $(Y, Y + Z)$:

$$\begin{aligned}\mathbf{P}(Y \in da, Y + Z \in db) &= \mathbf{P}(Y \in da, a + Z \in db) \\ &= \mathbf{P}(Y \in da) \mathbf{P}(a + Z \in db) \\ &= f(a) da g(b - a) db\end{aligned}$$

$$\mathbf{P}(Y \in da, Y + Z \in db) = f(a) g(b - a) da db$$

Integration über a gibt die Dichte von $Y + Z$:

$$\mathbf{P}(Y + Z \in db) = \left(\int f(a) g(b - a) da \right) db .$$

Beispiel: Für Y und Z unabhängig und $\text{Exp}(1)$ -verteilt ist

$$\begin{aligned} \mathbf{P}(Y + Z \in db) &= \left(\int_0^b e^{-a} e^{-(b-a)} da \right) db \\ &= b e^{-b} db, \quad b \geq 0. \end{aligned}$$

(Dichte der Gamma(2)-Verteilung)

3. Zerlegung der gemeinsamen Verteilung

(Buch S. 111)

Bisher legten wir das Hauptaugenmerk auf den

Aufbau der gemeinsamen Verteilung von X_1 und X_2

aus der Verteilung ρ von X_1

und Übergangswahrscheinlichkeiten $P(a_1, \cdot)$:

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) := \rho(a_1)P(a_1, a_2)$$

Jetzt:

Zerlegung der gemeinsamen Verteilung von X_1 und X_2

in die Verteilung von X_1

und die *bedingte Verteilung* von X_2 gegeben X_1

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \mathbf{P}(X_1 = a_1) \frac{\mathbf{P}(X_1 = a_1, X_2 = a_2)}{\mathbf{P}(X_1 = a_1)}$$

Sei X_1 eine diskrete Zufallsvariable mit Zielbereich S_1
und X_2 eine Zufallsvariable mit Zielbereich S_2 .

Dann ist die

bedingte Wahrscheinlichkeit von $\{X_2 \in A_2\}$,

gegeben $\{X_1 = a_1\}$

definiert als

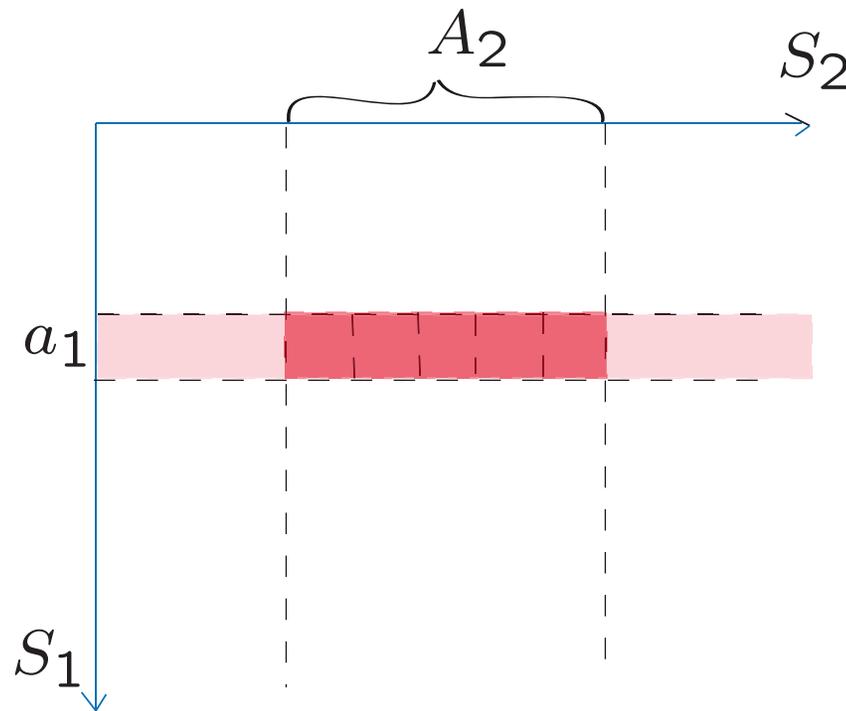
$$\mathbf{P}(X_2 \in A_2 \mid X_1 = a_1) := \frac{\mathbf{P}(X_1 = a_1, X_2 \in A_2)}{\mathbf{P}(X_1 = a_1)} .$$

In der Matrix der gemeinsamen Verteilungsgewichte

$$\mu(a_1, a_2) = \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

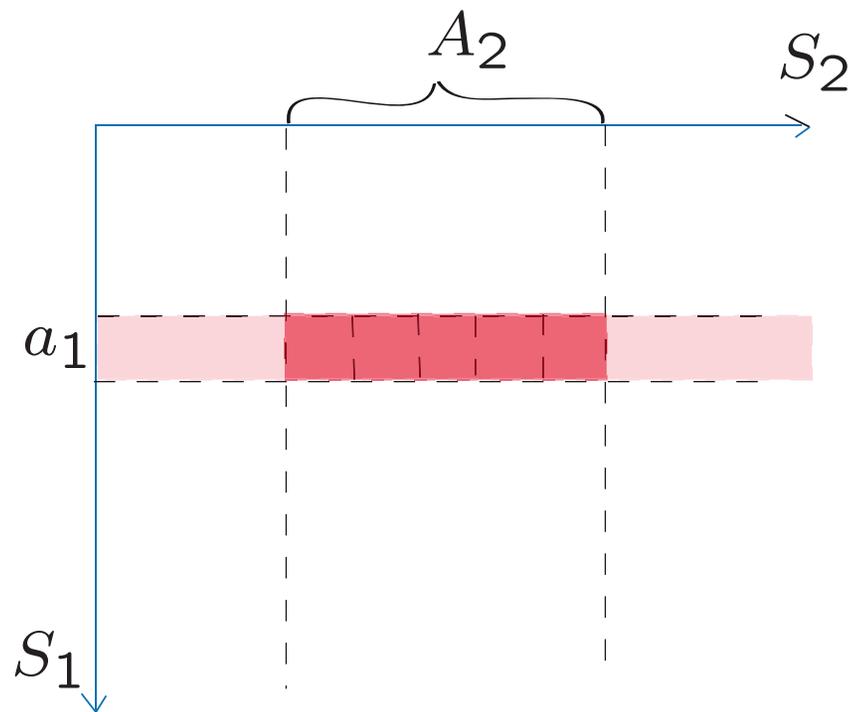
ist $\mathbf{P}(X_2 \in A_2 | X_1 = a_1)$ das relative Gewicht von A_2

bezogen auf das **Gesamtgewicht der Zeile a_1**



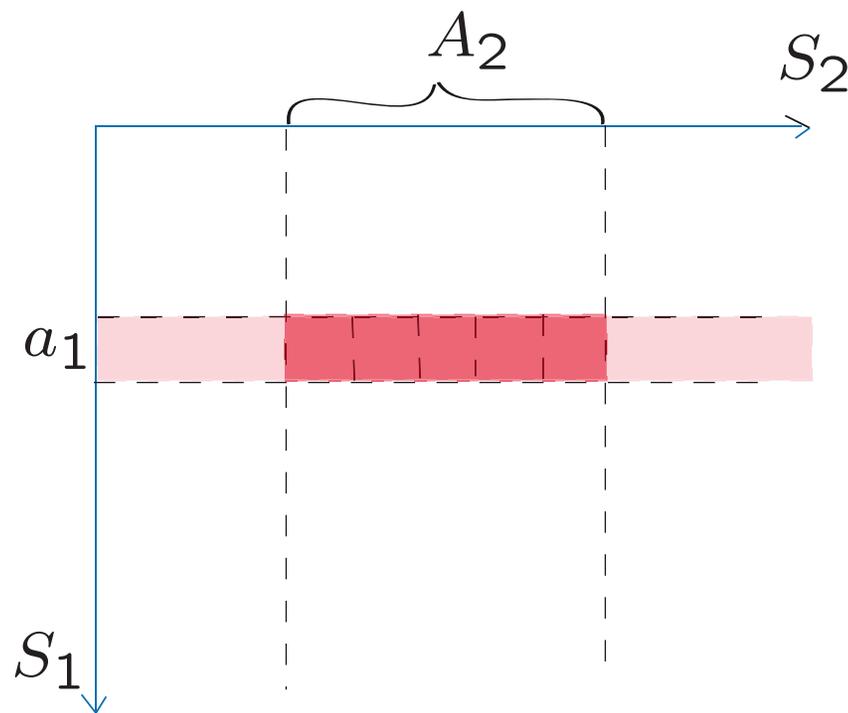
Die Verteilung $\mathbf{P}(X_2 \in \cdot | X_1 = a_1)$

heißt die *bedingte Verteilung* von X_2 , gegeben $\{X_1 = a_1\}$.



Definieren wir Übergangswahrscheinlichkeiten durch

$$P_{a_1}(X_2 \in A_2) := P(X_2 \in A_2 \mid X_1 = a_1)$$



Definieren wir Übergangswahrscheinlichkeiten durch

$$\mathbf{P}_{a_1}(X_2 \in A_2) := \mathbf{P}(X_2 \in A_2 \mid X_1 = a_1)$$

dann bekommen wir die
aus den vorigen Vorlesungen vertraute Formel

$$\mathbf{P}(X_1 = a_1, X_2 \in A_2) = \mathbf{P}(X_1 = a_1) \mathbf{P}_{a_1}(X_2 \in A_2).$$

4. “Wie war der erste Schritt?”

(Buch S. 111-112)

Bei der Untersuchung von zwei Zufallsvariablen X_1, X_2
kann man immer
zu einer **2-stufigen Betrachtungsweise** übergehen.

Man kann dabei wählen,
ob man X_1 in die erste Stufe aufnimmt oder in die zweite.

Beispiel:

Es seien Y und Z unabhängige \mathbb{Z} -wertige Zufallsvariable und

$$X_1 := Y, X_2 := Y + Z.$$

Wir haben gesehen:

Die bedingte Verteilung von $Y + Z$, gegeben $\{Y = a\}$,
ist die Verteilung von $a + Z$.

Was ergibt sich für die bedingte Verteilung von Y ,
gegeben $\{Y + Z = b\}$?

“Wie war der erste Schritt?”

Die bedingte Verteilung von Y , gegeben $Y + Z = b$,
hat die Gewichte

$$\mathbf{P}(Y = a \mid Y + Z = b) = \frac{\mathbf{P}(Y = a) \mathbf{P}(Z = b - a)}{\mathbf{P}(Y + Z = b)} .$$

Ein instruktiver Spezialfall:

Y und Z seien unabhängig und $\text{Geom}(p)$ -verteilt. Dann ist

$$\mathbf{P}(Y = a) \mathbf{P}(Z = b - a) = q^{a-1} p q^{(b-a)-1} p = p^2 q^{b-2} .$$

Dieses hängt nicht von a ab.

Also ist die bedingte Verteilung von Y gegeben $\{Y + Z = b\}$
die uniforme Verteilung auf $\{1, \dots, b - 1\}$

Das ist auch ohne Rechnung plausibel:

Gegeben, dass in einem p -Münzwurf
der zweite Erfolg beim b -ten Versuch kommt,
ist der Zeitpunkt des ersten Erfolges

uniform verteilt auf $\{1, \dots, b - 1\}$.

5. “Wann kamen die erfolgreichen Würfe”?

Beispiel: Erfolgreiche Würfe beim Münzwurf

Bei einem 10-maligen p -Münzwurf sei

K die Anzahl der Erfolge,

und $G \subset \{1, \dots, 10\}$ die zufällige Menge der Zeiten,
zu denen die Erfolge eintreten.

Wie ist die bedingte Verteilung von G , gegeben $\{K = 4\}$?

Für jede 4-elementige Teilmenge a von $\{1, \dots, 10\}$ ist

$$\mathbf{P}(G = a, K = 4) = \mathbf{P}(G = a) = p^4(1 - p)^6.$$

Das hängt nicht von a ab, also ist
die bedingte Verteilung von G uniform.

6. Bedingte Dichten

(Buch S. 112)

Ist $f(a_1, a_2) da_1 da_2$ gemeinsame Dichte von X_1 und X_2
und $f_1(a_1) da_1$ Dichte von X_1 ,

dann setzen wir

$$\mathbf{P}(X_2 \in da_2 \mid X_1 = a_1) := \frac{f(a_1, a_2)}{f_1(a_1)} da_2$$

und sprechen von der

bedingten Dichte von X_2 , gegeben $\{X_1 = a_1\}$.

Beispiel: Exponentialverteilte Summanden

Y und Z seien unabhängig und $\text{Exp}(1)$ -verteilt.

Was ist die bedingte Dichte von Y , gegeben $\{Y + Z = b\}$?

Die gemeinsame Dichte von Y und $Y + Z$ ist

$$e^{-a} e^{-(b-a)} da db = e^{-b} da db, \quad 0 \leq a \leq b$$

Die Dichte von $Y + Z$ ist $\left(\int_0^b da \right) e^{-b} db = b e^{-b} db$

Also:

$$\mathbf{P}(Y \in da | Y + Z = b) = \frac{1}{b e^{-b}} e^{-b} da = \frac{1}{b} da, \quad 0 \leq a \leq b.$$

7. Bedingter Erwartungswert

In einem zweistufigen Experiment hatten wir

$$\mathbf{E}_{a_1}[h(X_1, X_2)] = \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}_{a_1}(X_2 = a_2)$$

Wegen

$$\mathbf{P}_{a_1}(X_2 = a_2) = \mathbf{P}(X_2 = a_2 \mid X_1 = a_1)$$

ist damit die folgende Definition konsistent:

*Bedingter Erwartungswert von $h(X_1, X_2)$,
gegeben $\{X_1 = a_1\}$:*

$$\begin{aligned} & \mathbf{E}[h(X_1, X_2) \mid X_1 = a_1] \\ := & \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}(X_2 = a_2 \mid X_1 = a_1) \end{aligned}$$

Bedingte Erwartung von $h(X_1, X_2)$, gegeben X_1 :

$$\begin{aligned} & \mathbf{E}[h(X_1, X_2) \mid X_1] := e(X_1), \\ & \text{mit} \\ & e(a_1) := \mathbf{E}[h(X_1, X_2) \mid X_1 = a_1]. \end{aligned}$$

Zum Merken:

Der **bedingte Erwartungswert** von Y , gegeben $X = a$

(Symbol : $\mathbf{E}[Y \mid X = a]$ oder $\mathbf{E}_a[Y]$)

ist *der Erwartungswert unter der bedingten Verteilung*.

Im diskreten Fall

$$\sum_b b \mathbf{P}(Y = b \mid X = a),$$

und im Fall von Dichten

$$\int b \frac{f(a, b)}{f_1(a)} db.$$

Beispiel:

Z_1, \dots, Z_{10} sei ein p -Münzwurf der Länge 10, $K := \sum_{i=1}^{10} Z_i$.

Die *Runs* in (z_1, \dots, z_n) sind die
(in keinem größeren Block enthaltenen)

Blöcke aus nur Nullen oder nur Einsen.

Z. B. hat $(0, 1, 1, 0, 0, 1, 1, 0, 0, 0)$ fünf Runs:

0, 11, 00, 11, 000.

Sei R die Anzahl der Runs in (Z_1, \dots, Z_{10}) .

Gefragt ist nach $\mathbf{E}[R|K = 4]$.

$$\begin{aligned}
 R &= \sum_{i=1}^n I_{\{\text{beim } i\text{-ten Wurf beginnt ein Run}\}} \\
 &= 1 + \sum_{i=1}^9 I_{\{Z_i \neq Z_{i+1}\}}.
 \end{aligned}$$

Und die bedingte Verteilung von (Z_1, \dots, Z_{10}) gegeben $K = 4$ entsteht so, dass man aus den Plätzen $1, \dots, 10$ rein zufällig 4 auswählt, auf die man die 4 Einsen setzt.

$$\begin{aligned}
 &P(Z_i \neq Z_{i+1} | K = 4) = 2 \cdot \frac{4}{10} \cdot \frac{6}{9} \quad (\text{warum?}), \\
 \text{also ist der gesuchte Wert gleich} & \quad 1 + 9 \cdot 2 \cdot \frac{4}{10} \cdot \frac{6}{9} = \frac{29}{5}.
 \end{aligned}$$

8. Bedingte Wahrscheinlichkeiten

(Buch S. 115-117)

Definition.

Seien E_1, E_2 Ereignisse. Dann ist die *bedingte Wahrscheinlichkeit von E_2 , gegeben E_1* , definiert als

$$\mathbf{P}(E_2 | E_1) := \frac{\mathbf{P}(E_2 \cap E_1)}{\mathbf{P}(E_1)} = \mathbf{P}(I_{E_2} = 1 | I_{E_1} = 1)$$

... die Wahrscheinlichkeit von E_2 , wenn man schon weiß, dass E_1 eingetreten ist.

Eine vertraute Regel (für zweistufige Experimente)
im neuen Gewand:

Formel für die totale Wahrscheinlichkeit:

$$\mathbf{P}(X_2 \in A_2) = \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1) \mathbf{P}(X_2 \in A_2 \mid X_1 = a_1)$$

Zweistufigkeit - Spieß umgedreht:

$$\mathbf{P}(X_1 = a_1 | X_2 = a_2) = \frac{\mathbf{P}(X_2=a_2 | X_1=a_1)\mathbf{P}(X_1=a_1)}{\mathbf{P}(X_2=a_2)}$$

Formel von Bayes:

$$\mathbf{P}(X_1 = a_1 | X_2 = a_2) = \frac{\mathbf{P}(X_2=a_2 | X_1=a_1)\mathbf{P}(X_1=a_1)}{\sum_{b \in S_1} \mathbf{P}(X_2=a_2 | X_1=b)\mathbf{P}(X_1=b)}$$

$$\mathbf{P}(E_1 | E_2) = \frac{\mathbf{P}(E_2 | E_1)\mathbf{P}(E_1)}{\mathbf{P}(E_2 | E_1)\mathbf{P}(E_1) + \mathbf{P}(E_2 | E_1^c)\mathbf{P}(E_1^c)}$$

Beispiel: Bei einer bestimmten Reihenuntersuchung wird eine kranke Person in 100% der Fälle positiv getestet, eine gesunde Person in 1%.

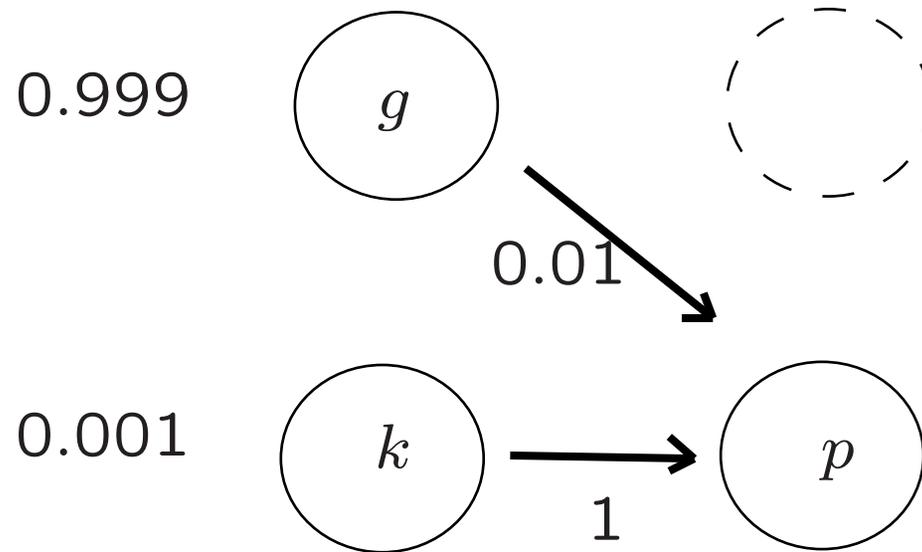
Wie groß ist die Wahrscheinlichkeit, dass eine positiv getestete Person wirklich krank ist?

Der Prozentsatz der kranken Personen sei 0.1% .

X_1 sei der Gesundheitszustand ($S_1 = \{g, k\}$),

X_2 der Testbefund ($S_2 = \{p, n\}$)

(X_1, X_2) entsteht über ein zweistufiges Experiment:



$$\mathbf{P}(X_1 = k | X_2 = p) = \frac{\mathbf{P}(X_1 = k, X_2 = p)}{\mathbf{P}(X_2 = p)}$$

$$= \frac{0.001 \cdot 1}{0.999 \cdot 0.01 + 0.001 \cdot 1} \approx \frac{1}{11}$$

Hier ist ein Rezept für eine intuitive Überschlagsrechnung,
gut gesehen von unserem Kameramann Sebastian Pohl:

In einer Population von 1000
sind 999 gesund und einer krank. Von den 999 Gesunden
werden ca 10 (falsch) positiv diagnostiziert.

Von 11 positiv Diagnostizierten
ist also im Schnitt nur einer krank.

9. Gedächtnislosigkeit der geometrischen und der Exponentialverteilung

(Buch S. 116)

Gedächtnislosigkeit der geometrischen Verteilung:

T sei $\text{Geom}(p)$ -verteilt. Dann gilt

$$\mathbf{P}(T > k + l \mid T > k) = q^{k+l} / q^k = q^l .$$

Die bedingte Verteilung von $T - k$, gegeben $\{T > k\}$,
ist somit gleich $\text{Geom}(p)$.

Die Kenntnis, dass T einen Wert größer als k annimmt,
ändert also die Verteilung für die “restliche Wartezeit” nicht.

Gedächtnislosigkeit der Exponentialverteilung

Für exponentialverteiltes T zum Parameter λ gilt für $r, s > 0$

$$\mathbf{P}(T > r + s \mid T > r) = e^{-\lambda s}.$$

Die bedingte Verteilung von $T - s$, gegeben $\{T > s\}$,
ist somit gleich $\text{Exp}(\lambda)$.