EINE WAHRSCHEINLICHKEITSVERTEILUNG OHNE GEDÄCHTNIS

von **D. J. Colwell** und **J. R. Gillett**
Staffordshire Polytechnic, Stafford, England

Originaltitel:
A Probability Distribution With No Memory

Aus: Teaching Statistics, Vol 12 No 3, Autumn 1990

Übertragen von Ingeborg Strauß, Kronberg im Taunus

Zusammenfassung:

Der folgende Artikel diskutiert und illustriert ein Paradoxon, das bei der Exponential-Verteilung auftritt. Bekannt ist dieses Phänomen als "Verlorenes Gedächtnis".

Die Poisson-Verteilung beschreibt das Verhalten einer diskreten Zufalls-Variablen. Sie resultiert aus einem Poisson-Prozeß. Schülern von Mathematik-Leistungskursen sollte diese wichtige Verteilung nicht vorenthalten werden. Beispiele für solche Situationen sind die Torschüsse beim Fußball, die auftretenden Fehler in einem Schreibmaschinentext, die Brenndauern von Glühbirnen und die Zeiten zwischen vorbeifahrenden Pkw's, gemessen von einem fixen Standpunkt aus am Autobahnrand.

Eine andere, der Poisson-Verteilung sehr ähnliche, ist die Exponential-Verteilung. Im Gegensatz zur Poisson-Verteilung ist sie stetig, wodurch die Intervalle zwischen den Ereignissen in einem Poisson-Prozeß ausgefüllt werden.

Eine Exponential-Verteilung mit dem Parameter $\lambda (>0)$ bzgl. einer Zufalls-Variablen X liegt genau dann vor, wenn die Dichtefunktion lautet:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Der Parameter ist ein Maß für die Rate, mit der die Ereignisse in dem Prozeß auftreten.

Es kann leicht gezeigt werden, daß gilt:

$E(X) = 1/\lambda$, $V(X) = 1/\lambda^2$ und $F(x) = P(X \le x) = 1 - e^{-\lambda x}$, $x \ge 0$.

Die Exponential-Verteilung hat eine besonders interessante Eigenschaft, die nicht sofort ins Auge springt. Angenommen, die Lebensdauer einer elektrischen Glühbirne werde durch eine Exponential-Verteilung X mit Parameter λ beschrieben. Wenn die Birne zu einem Zeitpunkt t_o noch nicht ausgebrannt ist, fragt man sich, wie die Verteilung bzgl. der verbleibenden Lebensdauer Y wohl aussieht?

Die Antwort ergibt sich, für $t>0$, aus folgenden Überlegungen:

$$
\begin{aligned}
P(Y>t) &= P(X>t_o+t \mid X>t_o) \\
&= \frac{P(X>t_o+t \text{ und } X>t_o)}{P(X>t_o)} \\
&= \frac{P(X>t_o+t)}{P(X>t_o)} \\
&= \frac{1 - P(X \le t_o+t)}{1 - P(X \le t_o)} \\
&= \frac{e^{-\lambda(t_o+t)}}{e^{-\lambda t_o}} \\
&= e^{-\lambda t}.
\end{aligned}
$$

Also ist $P(Y \le t) = 1 - e^{-\lambda t}$, $t \ge 0$, und es ist evident geworden, daß auch Y eine Exponential-Verteilung mit Parameter λ besitzt - dieselbe Verteilung wie für X.

Die Verteilung für die restliche Lebensdauer einer Glühbirne, die zum Zeitpunkt t_o noch ganz war, ist demnach dieselbe wie die für die gesamte Lebensdauer. Dieses offensichtlich paradoxe Resultat ist bekannt als das "Verlorene Gedächtnis" der Exponential-Verteilung.

Diese Eigenschaft kann sehr leicht durch Zuhilfenahme statistischer Software illustriert werden.

Man nehme zum Beispiel als mittlere Lebenszeit für eine Glühbirne 20 Zeiteinheiten. Abb. 1 zeigt das Histogramm mit 1000 Werten, das eine Exponential-Verteilung mit dem Mittelwert 20 (d.h. dem Parameter λ=0.05) darstellt.

Abb. 1: Histogramm mit 1000 Zufallswerten einer Exponential-Verteilung mit Parameter 0.05. Der Mittelwert und die Standardabweichung sind resp. 19.9 und 20.939.
Jeder * repräsentiert 5 Beobachtungen. 11 Werte sind größer als das letzte aufgeführte Klassenintervall.

Midpoint	Count	
2.50	220	**
7.50	170	**********************************
12.50	146	******************************
17.50	110	**********************
22.50	79	****************
27.50	60	************
32.50	55	***********
37.50	31	*******
42.50	24	*****
47.50	26	******
52.50	11	***
57.50	12	***
62.50	15	***
67.50	8	**
72.50	7	**
77.50	5	*
82.50	5	*
87.50	5	*

Figure 1

Setzen wir nun t_o=10. Abb. 2 zeigt das Histogramm für diese um 10 Zeiteinheiten reduzierten Werte, wobei negative Werte ignoriert wurden.

Abb. 2: Histogramm aus Abb. 1, reduziert um 10 Zeiteinheiten und nicht-negativ. Es gibt 610 solcher Werte. Der Mittelwert und die Standardabweichung betragen resp. 19.743 und 20.572.
Jeder * repräsentiert 5 Beobachtungen. 9 Werte sind größer als das letzte aufgeführte Klassenintervall.

Midpoint	Count	
2.50	146	******************************
7.50	110	**********************
12.50	79	****************
17.50	60	************
22.50	55	***********
27.50	31	*******
32.50	24	*****
37.50	26	******
42.50	11	***
47.50	12	***
52.50	15	***
57.50	8	**
62.50	7	**
67.50	5	*
72.50	5	*
77.50	5	*
82.50	1	*
87.50	1	*

Figure 2

Negative Werte wurden ignoriert, weil sie auf solche Glühbirnen hinweisen, die in weniger als 10 Zeiteinheiten kaputtgegangen sind.

Beide Histogramme zeigen dieselben Charakteristiken, und beide Datenmengen haben denselben Mittelwert und dieselbe Standardabweichung (≈ 20). Das Histogramm von Abb. 2 stimmt überein mit Daten, die aus einer Exponential-Verteilung mit dem Parameter $\lambda = 0.05$ gewonnen wurden.