

5. Wichtige diskrete Verteilungen

Wir diskutieren nun einige wichtige *diskrete* Verteilungen. Bei diesen Verteilungen handelt es sich um Funktionen, die von gewissen *Parametern* abhängen. Eigentlich betrachten wir also immer eine ganze Familie von ähnlichen Verteilungen.

5.1 Bernoulli-Verteilung

Eine Zufallsvariable X mit $W_X = \{0, 1\}$ und der Dichte

$$f_X(x) = \begin{cases} p & \text{für } x = 1, \\ 1 - p & \text{für } x = 0. \end{cases}$$

heißt **Bernoulli-verteilt**. Den Parameter p nennen wir **Erfolgswahrscheinlichkeit**.

Eine solche Verteilung erhält man z.B. bei einer einzelnen Indikatorvariablen. Es gilt mit $q := 1 - p$

$$\mathbb{E}[X] = p \text{ und } \text{Var}[X] = pq,$$

wegen $\mathbb{E}[X^2] = p$ und $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2$.

Der Name der Bernoulli-Verteilung geht zurück auf den Schweizer Mathematiker **Jakob Bernoulli** (1654–1705). Wie viele andere Mathematiker seiner Zeit hätte auch Bernoulli nach dem Wunsch seines Vaters ursprünglich Theologe werden sollen. Sein Werk *ars conjectandi* stellt eine der ersten Arbeiten dar, die sich mit dem Teil der Mathematik beschäftigen, den wir heute als Wahrscheinlichkeitstheorie bezeichnen.

5.2 Binomialverteilung

Eine Bernoulli-verteilte Zufallsvariable entspricht der Verteilung *einer* Indikatorvariablen. Häufig betrachtet man jedoch Summen von Indikatorvariablen.

Definition 54

Sei $X := X_1 + \dots + X_n$ als Summe von n unabhängigen, Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p definiert. Dann heißt X **binomialverteilt** mit den Parametern n und p . In Zeichen schreiben wir

$$X \sim \text{Bin}(n, p).$$

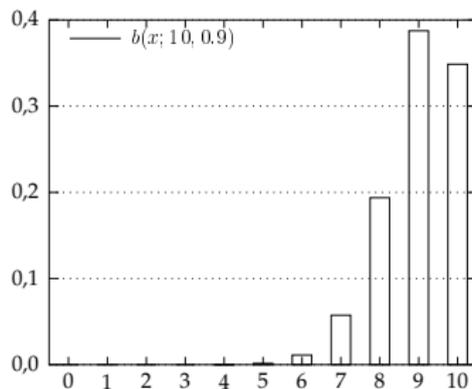
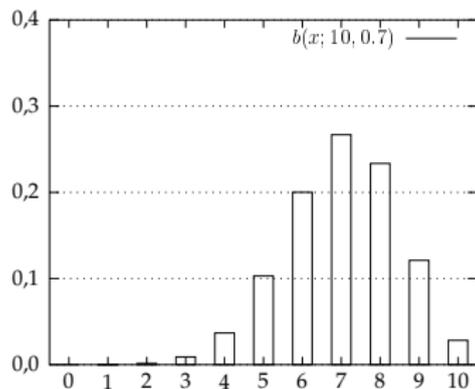
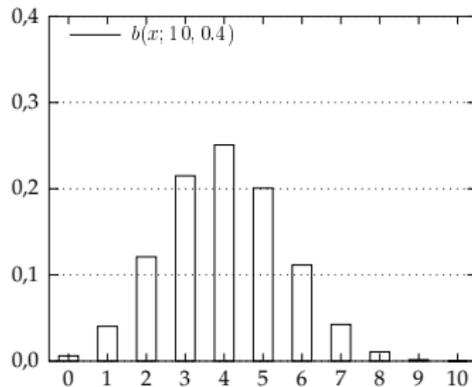
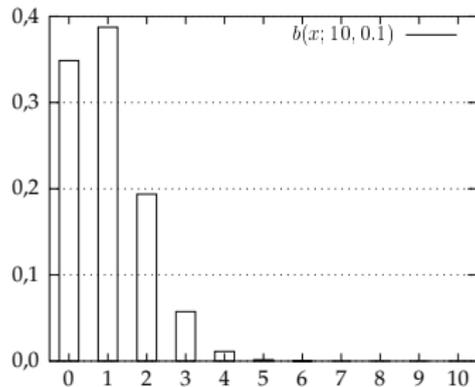
Es gilt $W_X = \{0, \dots, n\}$. Die Binomialverteilung besitzt die Dichte

$$f_X(x) := b(x; n, p) = \binom{n}{x} p^x q^{n-x}$$

mit $q := 1 - p$. Da die Binomialverteilung eine sehr wichtige Rolle spielt, führen wir für die Dichtefunktion die Abkürzung $b(x; n, p)$ ein.

Mit den Sätzen über Erwartungswert und Varianz von Summen unabhängiger Zufallsvariablen erhalten wir sofort

$$\mathbb{E}[X] = np \quad \text{und} \quad \text{Var}[X] = npq.$$



Dichte der Binomialverteilung

Satz 55

Wenn $X \sim \text{Bin}(n_x, p)$ und $Y \sim \text{Bin}(n_y, p)$ unabhängig sind, dann gilt für $Z := X + Y$, dass $Z \sim \text{Bin}(n_x + n_y, p)$.

Beweis:

Die Aussage folgt sofort, wenn man gemäß der Definition der Binomialverteilung X und Y als Summen von Indikatorvariablen darstellt. Z ist dann offensichtlich wieder eine Summe von unabhängigen Indikatorvariablen. □

5.3 Geometrische Verteilung

Definition 56

Die Dichte der **geometrischen Verteilung** mit Parameter/Erfolgswahrscheinlichkeit $p \in [0, 1]$ und $q := 1 - p$ ist gegeben durch

$$f_X(i) = pq^{i-1} \quad \text{für } i \in \mathbb{N}.$$

Für Erwartungswert und Varianz geometrisch verteilter Zufallsvariablen gilt

$$\mathbb{E}[X] = \frac{1}{p} \quad \text{und} \quad \text{Var}[X] = \frac{q}{p^2},$$

denn es gilt:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} i \cdot pq^{i-1} = p \cdot \sum_{i=1}^{\infty} i \cdot q^{i-1} = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}.$$

$\mathbb{E}[X^2]$ ergibt sich gemäß der Formel (siehe DS I)

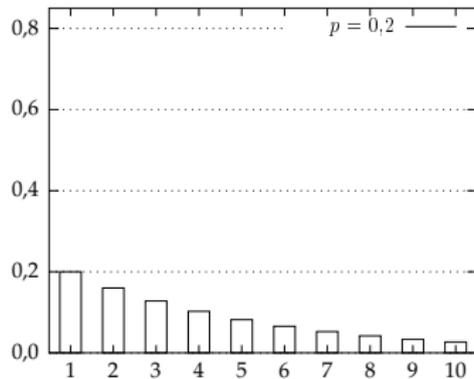
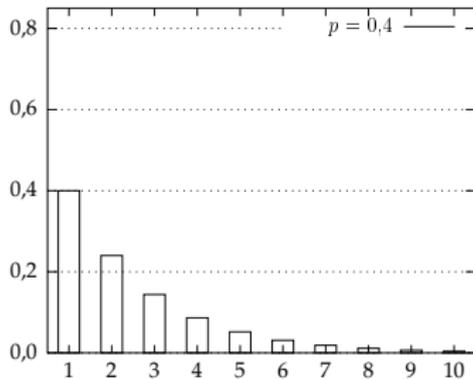
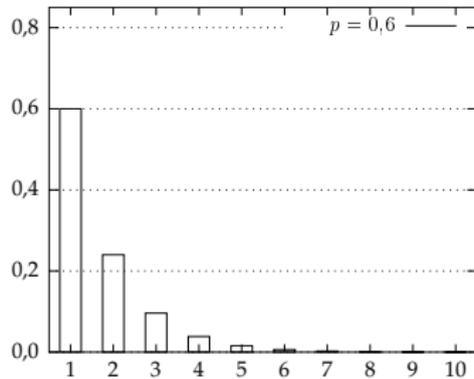
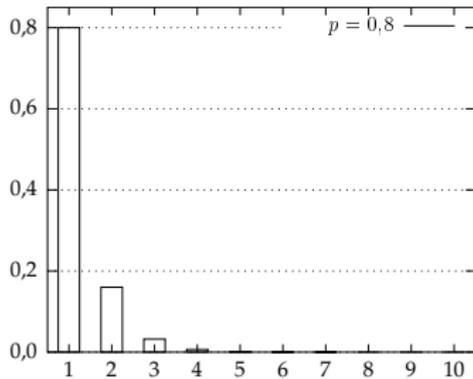
$$\sum_{n \geq 0} \binom{c+n-1}{n} z^n = \frac{1}{(1-z)^c} = (1-z)^{-c}$$

zu

$$\begin{aligned} \mathbb{E}[X^2] &= \sum_{i=1}^{\infty} i^2 \cdot pq^{i-1} \\ &= p \cdot \left(q \sum_{i=0}^{\infty} (i+2)(i+1) \cdot q^i + \sum_{i=0}^{\infty} (i+1) \cdot q^i \right) \\ &= \frac{q \cdot 2}{p^2} + \frac{1}{p} = \frac{2-p}{p^2}, \end{aligned}$$

und damit

$$\text{Var}[X] = \frac{q}{p^2}.$$



Dichte der geometrischen Verteilung

Sei X geometrisch verteilt mit Erfolgswahrscheinlichkeit p . Dann ist $\Pr[X = k]$ die Wahrscheinlichkeit, dass wir bei einem binären Experiment mit Erfolgswahrscheinlichkeit p genau in der k -ten unabhängigen Wiederholung das erste Mal erfolgreich sind.

Wie groß ist die Wahrscheinlichkeit $\Pr[X > y + x \mid X > x]$?

Da bei den ersten x Versuchen kein Erfolg eintrat, stellen wir uns vor, dass das „eigentliche“ Experiment erst ab dem $(x + 1)$ -ten Versuch beginnt. Die Zeit bis zum ersten Erfolg bei diesem neuen Experiment nennen wir X' . Damit $X > y + x$ gilt, muss $X' > y$ gelten. Es ist intuitiv, dass X' wieder geometrisch verteilt ist mit Erfolgswahrscheinlichkeit p , dass also für $x, y \in \mathbb{N}$ gilt:

$$\Pr[X > y + x \mid X > x] = \Pr[X' > y]. \quad (6)$$

Formal gilt

$$\begin{aligned}\Pr[X > x] &= \sum_{i=x+1}^{\infty} (1-p)^{i-1} p = (1-p)^x p \cdot \sum_{i=0}^{\infty} (1-p)^i \\ &= (1-p)^x p \cdot \frac{1}{1-(1-p)} = (1-p)^x,\end{aligned}$$

sowie

$$\begin{aligned}\Pr[X > y+x \mid X > x] &= \frac{\Pr[X > y+x, X > x]}{\Pr[X > x]} \\ &= \frac{\Pr[X > y+x]}{\Pr[X > x]} \\ &= (1-p)^{y+x} \cdot (1-p)^{-x} = (1-p)^y \\ &= \Pr[X > y].\end{aligned}$$

Diese Eigenschaft nennt man *Gedächtnislosigkeit*, da eine geometrisch verteilte Zufallsvariable gewissermaßen vergisst, dass sie schon x Misserfolge hinter sich hat und sich deshalb zum Zeitpunkt $y + x$ genauso verhält wie ursprünglich zur Zeit y .

Warten auf den n -ten Erfolg.

Wir betrachten n unabhängige Zufallsvariablen X_1, \dots, X_n , die jeweils geometrisch verteilt sind mit Parameter p , und bestimmen die Dichte der Zufallsvariablen $Z := X_1 + \dots + X_n$. Damit bezeichnet Z also die Anzahl der Versuche bis zum n -ten erfolgreichen Experiment (einschließlich).

Falls $Z = z$ ist, so werden also genau n erfolgreiche und $z - n$ nicht erfolgreiche Experimente durchgeführt. Dafür gibt es genau $\binom{z-1}{n-1}$ Möglichkeiten, von denen jede mit Wahrscheinlichkeit $p^n(1-p)^{z-n}$ eintritt. Es gilt also

$$f_Z(z) = \binom{z-1}{n-1} \cdot p^n(1-p)^{z-n}.$$

Die Zufallsvariable Z nennt man **negativ binomialverteilt** mit Ordnung n .

Das Coupon-Collector-Problem

In manchen Branchen legen Firmen den Verpackungen ihrer Produkte oft kleine Bilder oder andere Gegenstände bei, um den Käufer zum Sammeln anzuregen. Wenn es insgesamt n verschiedene solche Beilagen gibt, wie viele Packungen muss man im Mittel erwerben, bis man eine vollständige Sammlung besitzt? Hierbei nehmen wir an, dass bei jedem Kauf jede Beilage mit gleicher Wahrscheinlichkeit auftritt.

Sei

- X die Anzahl der zu tätigen Käufe, und
- bezeichne Phase i die Schritte vom Erwerb der $(i - 1)$ -ten Beilage (ausschließlich) bis zum Erwerb der i -ten Beilage (einschließlich).

Sei etwa $n = 4$, und seien die Beilagen mit den Zahlen 1, 2, 3, 4 identifiziert. Ein Experiment ist z.B.:

$$\underbrace{2}_1, \underbrace{2, 1}_2, \underbrace{2, 2, 3}_3, \underbrace{1, 3, 2, 3, 1, 4}_4 .$$

Beobachtung:

Phase i endet genau dann, wenn wir eine der $n - i + 1$ Beilagen erhalten, die wir noch nicht besitzen.

Somit ist X_i geometrisch verteilt mit Parameter $p = \frac{n-i+1}{n}$ und es gilt $\mathbb{E}[X_i] = \frac{n}{n-i+1}$.

Damit folgt aber sofort

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=1}^n \mathbb{E}[X_i] \\ &= \sum_{i=1}^n \frac{n}{n-i+1} \\ &= n \cdot \sum_{i=1}^n \frac{1}{i} = n \cdot H_n,\end{aligned}$$

wobei $H_n := \sum_{i=1}^n \frac{1}{i}$ die n -te harmonische Zahl bezeichnet. Da $H_n = \ln n + O(1)$, folgt $\mathbb{E}[X] = n \ln n + O(n)$.

5.4 Poisson-Verteilung

Die Poisson-Verteilung mit dem Parameter λ hat den Wertebereich $W_X = \mathbb{N}_0$ und besitzt die Dichte

$$f_X(i) = \frac{e^{-\lambda} \lambda^i}{i!} \quad \text{für } i \in \mathbb{N}_0.$$

f_X ist eine zulässige Dichte, da

$$\begin{aligned} \sum_{i=0}^{\infty} f_X(i) &= \sum_{i=0}^{\infty} \frac{e^{-\lambda} \lambda^i}{i!} \\ &= e^{-\lambda} \cdot e^{\lambda} = 1. \end{aligned}$$

Für den Erwartungswert erhalten wir

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda} \lambda^i}{i!} \\ &= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} \\ &= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \\ &= \lambda e^{-\lambda} e^{\lambda} = \lambda.\end{aligned}$$

Da

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{i=0}^{\infty} i(i-1) \cdot \frac{e^{-\lambda} \lambda^i}{i!} \\ &= \lambda^2 e^{-\lambda} \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} \\ &= \lambda^2 e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \\ &= \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2\end{aligned}$$

und

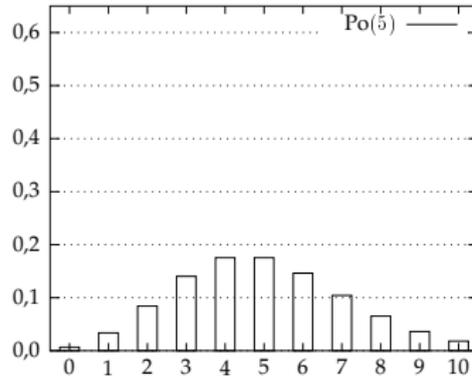
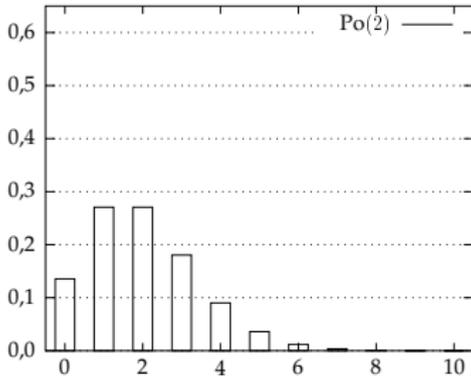
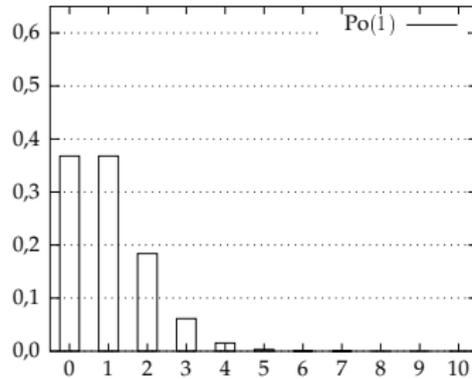
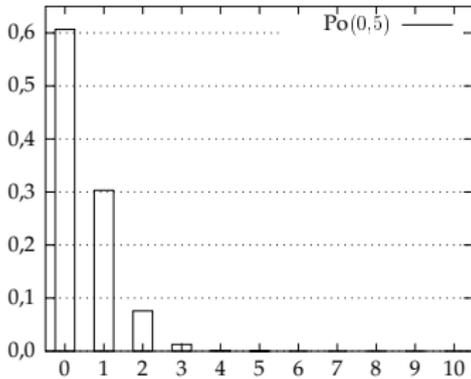
$$\begin{aligned}\mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\ = \mathbb{E}[X^2] - \mathbb{E}[X] + \mathbb{E}[X] - \mathbb{E}[X]^2 = \text{Var}[X],\end{aligned}$$

folgt

$$\text{Var}[X] = \mathbb{E}[X(X - 1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \quad (7)$$

Dafür, dass eine Zufallsvariable X Poisson-verteilt mit Parameter λ ist, schreiben wir auch

$$X \sim \text{Po}(\lambda).$$



Dichte der Poisson-Verteilung

5.4.1 Poisson-Verteilung als Grenzwert der Binomialverteilung

Wir betrachten eine Folge von binomialverteilten Zufallsvariablen X_n mit $X_n \sim \text{Bin}(n, p_n)$, wobei $p_n = \lambda/n$. Für ein beliebiges k mit $0 \leq k \leq n$ ist die Wahrscheinlichkeit, dass X_n den Wert k annimmt, gleich

$$\begin{aligned} b(k; n, p_n) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{(n \cdot p_n)^k}{k!} \cdot \frac{n^{\underline{k}}}{n^k} \cdot (1 - p_n)^{-k} \cdot (1 - p_n)^n \\ &= \frac{\lambda^k}{k!} \cdot \frac{n^{\underline{k}}}{n^k} \cdot \left(1 - \frac{\lambda}{n}\right)^{-k} \cdot \left(1 - \frac{\lambda}{n}\right)^n. \end{aligned}$$

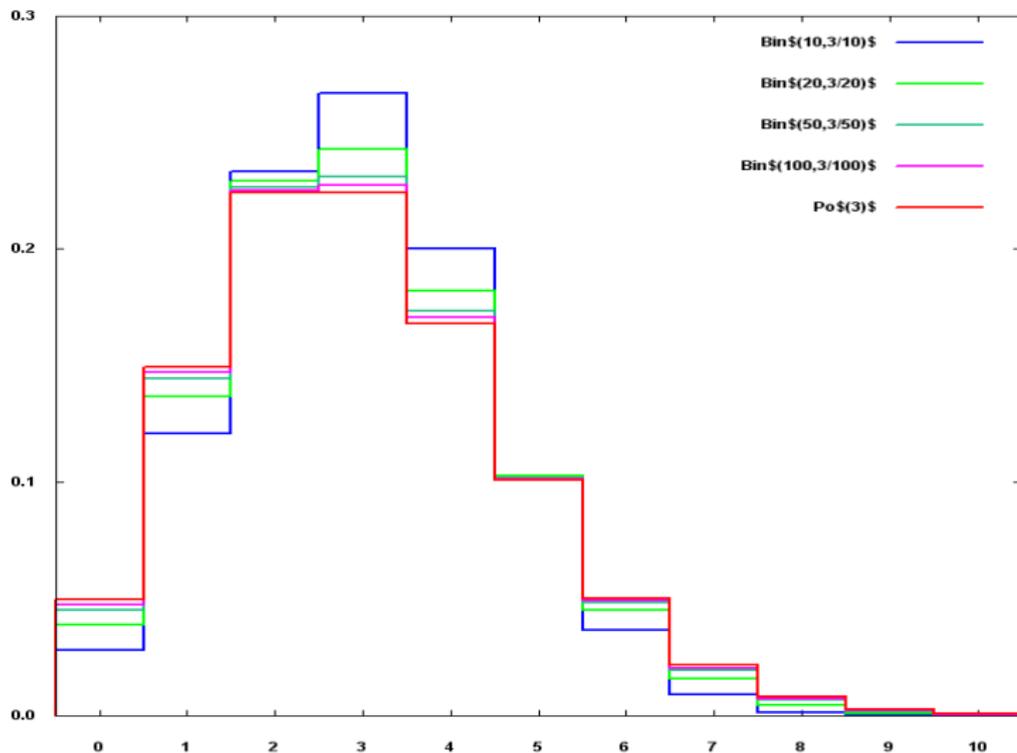
Wir betrachten nun $n \rightarrow \infty$ und erinnern uns, dass

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{n^k}{n^k} &= 1, \\ \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-k} &= 1, \text{ und} \\ \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n &= e^{-\lambda}.\end{aligned}$$

Damit folgt

$$\lim_{n \rightarrow \infty} b(k; n, p_n) = \lim_{n \rightarrow \infty} \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Die Wahrscheinlichkeit $b(k; n, p_n)$ konvergiert also für $n \rightarrow \infty$ gegen die Wahrscheinlichkeit, dass eine Poisson-verteilte Zufallsvariable mit Parameter λ den Wert k annimmt. Insgesamt folgt somit, dass die Verteilung einer Zufallsvariablen $X \sim \text{Bin}(n, \lambda/n)$ sich für $n \rightarrow \infty$ der Poisson-Verteilung $\text{Po}(\lambda)$ annähert.



Vergleich von Binomial- und Poisson-Verteilung

Ist also n im Vergleich zu λ hinreichend groß, so kann man die Poisson-Verteilung als Approximation der Binomialverteilung verwenden.

Diese Tatsache wird manchmal auch als **Gesetz seltener Ereignisse** bezeichnet, da die Wahrscheinlichkeit eines einzelnen Treffers $p_n = \lambda/n$ relativ klein sein muss, wenn die Approximation gute Ergebnisse liefern soll.

Die folgenden Voraussetzungen müssen erfüllt sein, damit die Annahme der Poisson-Verteilung gerechtfertigt ist:

- Die Ereignisse treten nie zur gleichen Zeit auf.
- Die Wahrscheinlichkeit, dass ein Ereignis in einem (kleinen) Zeitintervall δt auftritt, ist proportional zur Länge von δt .
- Die Anzahl der Ereignisse in einem festen Zeitintervall hängt nur von dessen Länge ab, nicht aber von der Lage auf der Zeitachse.
- Wenn man zwei disjunkte Zeitintervalle betrachtet, so sind die Anzahlen der Ereignisse in diesen Zeiträumen voneinander unabhängig.