7 Diskrete Verteilungen

Inhalt

7.1 Bernoulli-Verteilung

Definition 7.1 (Bernoulli-Experiment)
Ein ***Bernoulli-Experiment*** ist ein Zufallsexperiment, das nur zwei mögliche Ergebnisse hat: Ein bestimmtes Ereignis tritt ein oder nicht.

Beispiel 7.2 (Bernoulli-Experimente)
Beispiele für Bernoulli-Experimente:

1. Werfen einer Münze mit den Ergebnissen „Zahl“ oder „Wappen“
2. Werfen eines Würfels mit den Ergebnissen „Augenzahl 6“ oder „Augenzahl nicht 6“
3. Zufällige Auswahl eines Artikels mit den Ergebnissen „defekt“ oder „funktionstüchtig“

Definition 7.3 (Bernoulli-Verteilung)
Eine diskrete Zufallsvariable X heißt ***Bernoulli-verteilt***, wenn sie zwei Realisierungen $X \in \{0,1\}$ hat, deren Wahrscheinlichkeiten für ein $p \in [0,1]$ gegeben sind durch

$$P(X=k) = \begin{cases} p, & \text{falls} \quad k=1\,, \\ 1-p, & \text{falls} \quad k=0\,. \end{cases}$$

Kurzschreibweise: $X \sim \mathrm{B}(1,p)$.

Das „B“ in der Kurzschreibweise steht für die Binomialverteilung; der Zusammenhang wird im folgenden Abschnitt erläutert.

Welche Werte haben der Erwartungswert und die Varianz einer Bernoulli-verteilten Zufallsvariablen $X \sim \mathrm{B}(1, p)$?

$$\begin{aligned} \mathrm{E}(X) &= \sum_{i=0}^{1} x_i\, p_i = 0 \cdot (1-p) + 1 \cdot p = p\,, \\ \mathrm{Var}(X) &= \sum_{i=0}^{1} x_i^2\, p_i - (\mathrm{E}(X))^2 = 0 \cdot (1-p) + 1 \cdot p - p^2 = p\,(1-p)\,. \end{aligned}$$

Satz 7.4 (Erwartungswert und Varianz bei Bernoulli-Verteilung)
Für eine Bernoulli-verteilte Zufallsvariable $X \sim \mathrm{B}(1, p)$ gilt

$$\mathrm{E}(X) = p\,, \qquad \mathrm{Var}(X) = p\,(1-p)\,.$$

7.2 Binomialverteilung

Definition 7.5 (Bernoulli-Kette)
Die n-malige aufeinanderfolgende Ausführung eines Bernoulli-Experiments unter gleichen Bedingungen heißt ***Bernoulli-Kette der Länge*** n.

Beispiel : Bernoulli-Ketten

Beispiele für Bernoulli-Ketten (siehe Beispiel 7.2):

1. n-maliges Werfen einer Münze mit den Ergebnissen „Zahl“ oder „Wappen“
2. n-maliges Werfen eines Würfels mit den Ergebnissen „Augenzahl 6“ oder „Augenzahl nicht 6“
3. n-malige zufällige Auswahl mit Zurücklegen eines Artikels mit den Ergebnissen „defekt“ oder „funktionstüchtig“

Gegeben sei eine Bernoulli-Kette der Länge n, bei der in jeder Durchführung ein Ereignis mit der Wahrscheinlichkeit p eintritt. Sei X die diskrete Zufallsvariable, die der Anzahl der Durchführungen entspricht, in denen das Ereignis eintritt.

Im Beispiel zur verbogenen Münze des vorigen Kapitels haben wir hierfür die Wahrscheinlichkeitsverteilung

$$P(X = k) = \binom{n}{k} p^k\,(1-p)^{n-k}$$

erhalten.

Definition 7.6 (Binomialverteilung)
Sei $n \in \mathbb{N}$, und sei X eine diskrete Zufallsvariable mit den Realisierungen $X \in \{0,\dots,n\}$.

X heißt ***binomialverteilt***, wenn die Wahrscheinlichkeit $P(X=k)$ für $k=0,\dots,n$ und $p \in [0,1]$ gegeben ist durch

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Kurzschreibweise: $X \sim \mathrm{B}(n,p)$.

Der Spezialfall $\mathrm{B}(1,p)$ entspricht der Bernoulli-Verteilung.

Welche Werte haben der Erwartungswert und die Varianz einer binomialverteilten Zufallsvariable $X \sim \mathrm{B}(n,p)$?

Sei $X_i \in \{0,1\}$ das Ergebnis der i-ten Durchführung des Bernoulli-Experiments. Gemäß Satz 7.4 gilt für beliebiges i

$$\mathrm{E}(X_i) = p, \quad \mathrm{Var}(X_i) = p(1-p).$$

Für die Bernoulli-Kette folgt mit $X = X_1 + \ldots + X_n$

$$\begin{aligned} \mathrm{E}(X) &= \sum_{i=1}^{n} \mathrm{E}(X_i) = n\,\mathrm{E}(X_1) = n\,p, \\ \mathrm{Var}(X) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\,\mathrm{Var}(X_1) = n\,p(1-p). \end{aligned}$$

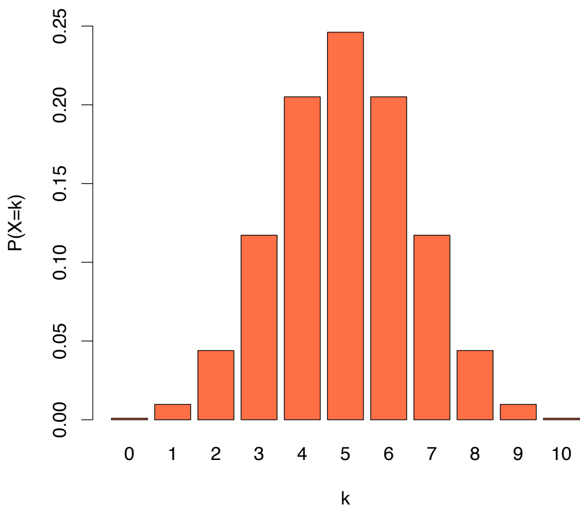
Die zweite Gleichung gilt, weil die Zufallsvariablen X_i bei Ausführung des Zufallsexperimentes unter gleichen Bedingungen stochastisch unabhängig sind.

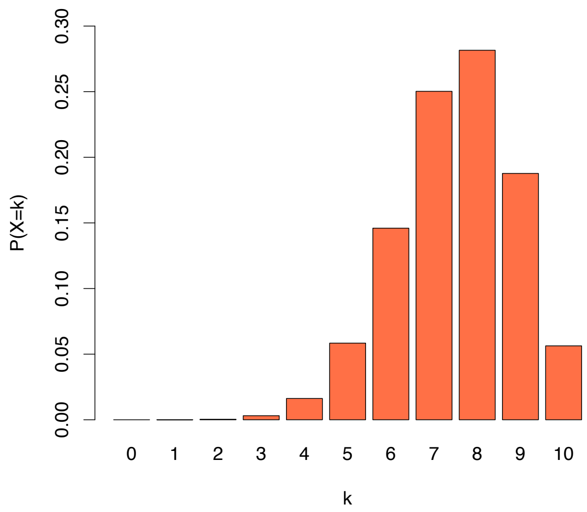
Satz 7.7 (Erwartungswert und Varianz bei Binomialverteilung)
Für eine binomialverteilte Zufallsvariable $X \sim \mathrm{B}(n,p)$ gilt

$$\mathrm{E}(X) = n\,p, \quad \mathrm{Var}(X) = n\,p\,(1-p).$$

gilt weil X_i stochastisch unabhängig sind

Beispiel : Verschiedene Binomialverteilungen
Links: $X \sim \mathrm{B}\left(10, \frac{1}{2}\right)$, $\mu = 5$, $\sigma = \sqrt{2.5} \approx 1.58$
Rechts: $X \sim \mathrm{B}\left(10, \frac{3}{4}\right)$, $\mu = 7.5$, $\sigma = \sqrt{\frac{15}{8}} \approx 1.37$





Satz 7.8 (Verteilungen abgeleiteter Zufallsvariablen)

(i) Sind die binomialverteilten Zufallsvariablen $X_1 \sim \mathrm{B}(n_1, p)$ und $X_2 \sim \mathrm{B}(n_2, p)$ stochastisch unabhängig, so gilt

$$X_1 + X_2 \sim \mathrm{B}(n_1 + n_2, p).$$

(ii) Für eine binomialverteilte Zufallsvariable $X \sim \mathrm{B}(n, p)$ ist $Y = n - X$ ebenfalls binomialverteilt mit

$$Y \sim \mathrm{B}(n, 1-p).$$

7.3 Hypergeometrische Verteilung

Gegeben sei eine Grundgesamtheit von N Elementen, von denen M eine bestimmte Eigenschaft haben. Aus der Menge werde eine Stichprobe vom Umfang n entnommen. Sei X die diskrete Zufallsvariable, die der Anzahl der Elemente in der Stichprobe mit der gewünschten Eigenschaft entspricht.

Im Beispiel zu den defekten USB-Sticks in Kapitel 4 haben wir mit Hilfe einer kombinatorischen Argumentation (Ziehen ohne Zurücklegen) die Wahrscheinlichkeitsverteilung

$$\boxed{P(X = m) = \frac{\binom{M}{m} \cdot \binom{N-M}{n-m}}{\binom{N}{n}}}$$

erhalten.

Definition 7.9 (Hypergeometrische Verteilung)
Sei $n \in \mathbb{N}$, und sei X eine diskrete Zufallsvariable mit den Realisierungen $X \in \{0,\dots,n\}$.

X heißt ***hypergeometrisch verteilt***, wenn die Wahrscheinlichkeit $P(X=k)$ für $k=0,\dots,n$ und Parameter $N,M,n \in \mathbb{N}$ mit $M \leq N$ und $n \leq N$ gegeben ist durch

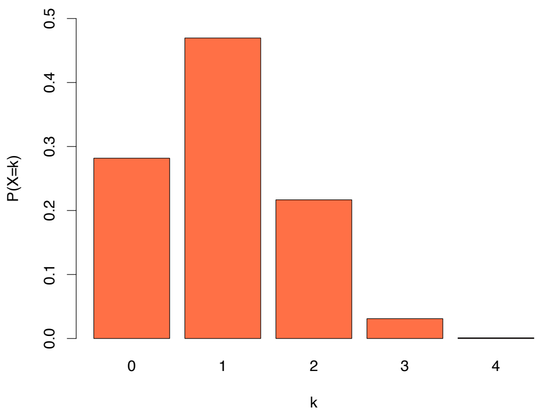
$$P(X=k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}.$$

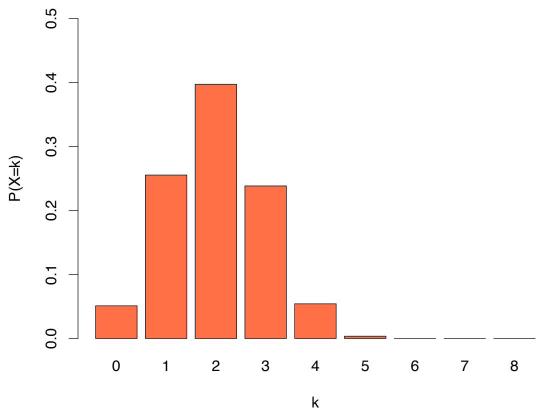
Kurzschreibweise: $X \sim \mathrm{Hyp}(N,M,n)$. – ohne Zurücklegen!

Satz 7.10 (Erwartungswert/Varianz bei hypergeometrischer Verteilung)
Für eine hypergeometrisch verteilte Zufallsvariable $X \sim \mathrm{Hyp}(N,M,n)$ gilt

$$\mathrm{E}(X) = n\frac{M}{N}, \quad \mathrm{Var}(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}.$$

Beispiel : Hypergeometrische Verteilungen
Links: $X \sim \mathrm{Hyp}(20,5,4)$, $\mu = 1$, $\sigma = \sqrt{\frac{12}{19}} \approx 0.795$
Rechts: $X \sim \mathrm{Hyp}(20,5,8)$, $\mu = 2$, $\sigma = \sqrt{\frac{18}{19}} \approx 0.973$





Stichproben vom Umfang 4 (links) bzw. 8 (rechts).

Ist der Umfang der Grundgesamtheit sehr viel größer als der Stichprobenumfang, $N \gg n$, so ändert sich durch das Ziehen der Stichprobe ohne Zurücklegen die Zusammensetzung der Grundgesamtheit nur wenig. Daher kann in diesem Fall an Stelle der hypergeometrischen Verteilung die rechnerisch einfachere Binomialverteilung verwendet werden.

Satz 7.11 (Faustregel für die hypergeometrische Verteilung)
Sei $X \sim \mathrm{Hyp}(N,M,n)$ eine hypergeometrisch verteilte Zufallsvariable.

Dann gilt die folgende Faustregel: Falls

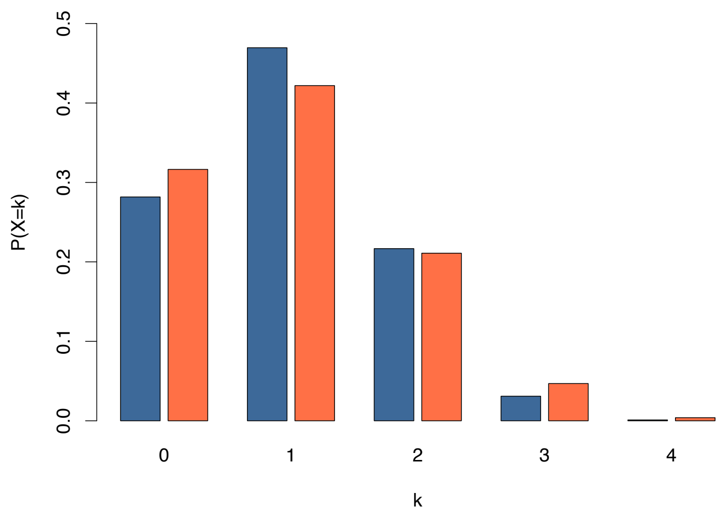
$$n \lesssim \frac{N}{20},$$

kann X näherungsweise durch die Binomialverteilung $\mathrm{B}(n,p)$ mit $p = M/N$ beschrieben werden.

Beispiel : Beispiel zur Faustregel für die hypergeometrische Verteilung

Blau: $X \sim \mathrm{Hyp}(20,5,4)\,, \quad \mu = 1\,, \quad \sigma = \sqrt{\frac{12}{19}} \approx 0.795$

Rot: $X \sim \mathrm{B}\left(4, \frac{5}{20}\right) = \mathrm{B}\left(4, \frac{1}{4}\right)\,, \quad \mu = 1\,, \quad \sigma = \sqrt{\frac{3}{4}} \approx 0.866$



Wegen $\quad 4 > \frac{20}{20} = 1 \quad$ ist die Näherung aus Satz 7.11 hier nicht gültig; die Verteilungen unterscheiden sich deutlich.

7.4 Poisson-Verteilung

Definition 7.12 (Poisson-Verteilung)
Sei X eine diskrete Zufallsvariable mit den abzählbar unendlich vielen Realisierungen $X \in \{0,1,\ldots\}$.

X heißt ***Poisson-verteilt***, wenn die Wahrscheinlichkeit $P(X = k)$ für $k = 0,1,\ldots$ und einen Parameter $\lambda \in [0,\infty)$ gegeben ist durch

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}\,.$$

Kurzschreibweise: $\quad X \sim \mathrm{Po}(\lambda)\,.$

Welchen Wert hat der Erwartungswert einer Poisson-verteilten Zufallsvariable $X \sim \mathrm{Po}(\lambda)$?

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} \\
&= \lambda\, e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda\, e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\
&= \lambda\, e^{-\lambda} e^{\lambda} = \lambda\,.
\end{aligned}$$

In der vorletzten Umformung entspricht die Summe der Potenzreihe der Exponentialfunktion. Die Varianz kann auf ähnliche Weise berechnet werden.

Satz 7.13 (Erwartungswert und Varianz bei der Poisson-Verteilung)
Für eine Poisson-verteilte Zufallsvariable $X \sim \mathrm{Po}(\lambda)$ gilt

$$\mathrm{E}(X) = \lambda\,, \qquad \mathrm{Var}(X) = \lambda\,.$$

Satz 7.14 (Zusammenhang von Poisson- und Binomialverteilung)
Sei $X \sim \mathrm{B}(n, p)$ eine binomialverteilte Zufallsvariable.

Für $n \to \infty$ und $p \to 0$ nähert sich die Wahrscheinlichkeitsverteilung von X an die Poisson-Verteilung an:

$$\lim_{\substack{n\to\infty \\ p\to 0}} P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{mit} \quad \lambda = n\,p\,.$$

Nach einer Faustregel ist die Poisson-Verteilung $\mathrm{Po}(n\,p)$ eine gute Näherung der Binomialverteilung $\mathrm{B}(n, p)$, wenn $n \geq 20$ und $p \leq 0.05$ sind.

Beispiel : Zusammenhang von Poisson- und Binomialverteilung

Eine Maschine produziere pro Tag 100 Artikel mit 5 % Ausschuss.

Mit welcher Wahrscheinlichkeit sind in einer Tagesproduktion 5 oder weniger Artikel defekt?

Wie hoch sind Erwartungswert und Varianz der Anzahl defekter Artikel?

1. Exakte Lösung mit Binomialverteilung:

Für die Anzahl der defekten Artikel X gilt

$$X \sim \mathrm{B}(n, p) \quad \text{mit} \quad n = 100 \quad \text{und} \quad p = 0.05$$

$$\Rightarrow \quad P(X \leq 5) = \sum_{k=0}^{5} P(X = k) = \sum_{k=0}^{5} \binom{n}{k} p^k (1-p)^{n-k} \approx 0.61606\,,$$

$$\mathrm{E}(X) = n\,p = 5\,, \quad \mathrm{Var}(X) = n\,p\,(1-p) = 4.75\,.$$

2. Näherungsweise Lösung mit Poisson-Verteilung:

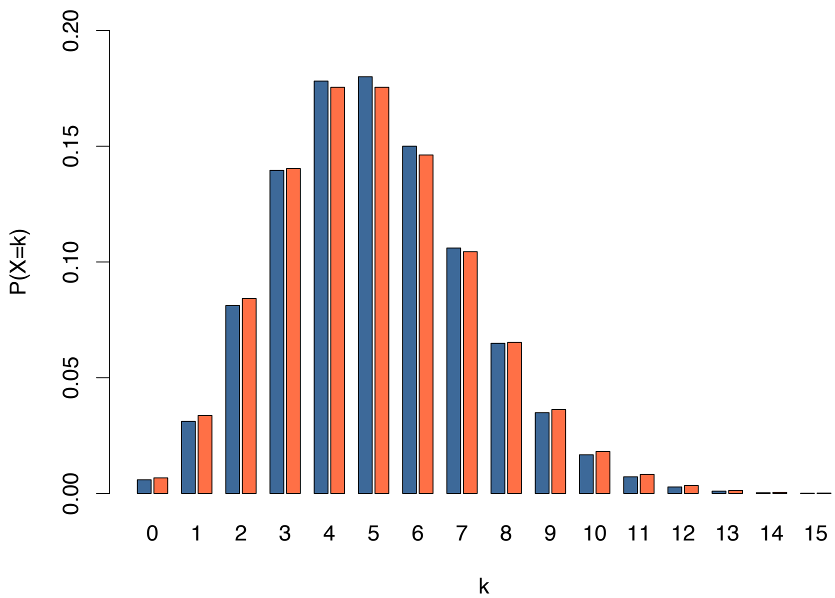
Für die näherungsweise Anzahl der defekten Artikel Y gilt

$$Y = \mathrm{Po}(\lambda) \quad \text{mit} \quad \lambda = n\,p = 5$$
$$\Rightarrow \quad P(Y \le 5) = \sum_{k=0}^{5} P(Y=k) = e^{-5} \sum_{k=0}^{5} \frac{5^k}{k!} \approx 0.61596\,.$$
$$\mathrm{E}(Y) = n\,p = 5\,, \quad \mathrm{Var}(X) = n\,p = 5\,.$$

Der Rechenaufwand für Lösungsansatz 2 ist geringer als für Ansatz 1; außerdem muss in Ansatz 2 nicht mit großen Fakultäten hantiert werden.

Blau: $X \sim \mathrm{B}(100, 0.05)\,, \quad \mu = 5\,, \quad \sigma = \sqrt{\frac{19}{4}} \approx 2.18$

Rot: $X \sim \mathrm{Po}(5)\,, \quad \mu = 5\,, \quad \sigma = \sqrt{5} \approx 2.24$



Die Poisson-Verteilung ist hier eine gute Näherung der Binomialverteilung.

7.5 Poisson-Prozess

Eine häufige Anwendung der Poisson-Verteilung ist die Modellierung seltener Ereignisse in einem bestimmten Zeitraum. Dies führt zum ***Poisson-Prozess***.

Definition 7.15 (Poisson-Prozess)
Gegeben sei eine Familie (X_t) von Zufallsvariablen X_t mit $t \in \mathbb{R}_0^+$.

Die Familie (X_t) heißt ***Poisson-Prozess***, falls alle X_t Poisson-verteilt sind mit

$$X_t \sim \mathrm{Po}(t\,\lambda)\,,$$

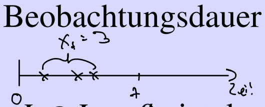
falls also gilt

$$P(X_t = k) = \frac{(t\,\lambda)^k}{k!}\, e^{-t\,\lambda}\,.$$

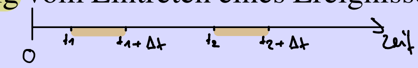
Satz 7.16 (Poisson-Prozess)
Sei (X_t) eine Familie von Zufallsvariablen, wobei X_t der Anzahl des Eintretens eines bestimmten Ereignisses im Zeitintervall $[0,t]$ entspricht.

(X_t) ist ein Poisson-Prozess, wenn folgende Bedingungen erfüllt sind.

(i) Die Wahrscheinlichkeit für das Eintreten eines Ereignisses ist proportional zur Beobachtungsdauer Δt und unabhängig davon, ab welchem Zeitpunkt t die Beobachtungsdauer beginnt.



(ii) Für disjunkte Zeitintervalle $I_1 = [t_1, t_1 + \Delta t]$ und $I_2 = [t_2, t_2 + \Delta t]$, d. h. $I_1 \cap I_2 = \emptyset$, ist das Eintreten eines Ereignisses in I_1 stochastisch unabhängig vom Eintreten eines Ereignisses in I_2.



(iii) Für $\Delta t \to 0$ ist die Wahrscheinlichkeit für das Eintreten von mehr als einem Ereignis in einem Zeitintervall der Länge Δt proportional zu $(\Delta t)^2$ und damit vernachlässigbar klein.
– zwei Ereignisse gleichzeitig?

Wegen der letzten Bedingung ist der Poisson-Prozess zur Modellierung (vergleichsweise) seltener Ereignisse geeignet, wie z. B.

- das Auftreten von Fehlern, Unfällen oder radioaktiven Zerfallsereignissen,
- das Eintreffen von Telefonanrufen oder Server-Anfragen,
- das Eintreffen von Kunden oder Artikeln in einer Warteschlange.

Der Parameter t muss nicht der Zeit entsprechen, er kann u. a. auch einen Ort oder eine Position darstellen. Beispielsweise könnte das Auftreten von Programmierfehlern in Code-Abschnitten der Länge t als Poisson-Prozess modelliert werden.

Satz 7.17 (Ereignisse in einem Zeitintervall)
Sei (X_t) ein Poisson-Prozess mit $X_t \sim \text{Po}(t\lambda)$.

Dann ist die Anzahl $X_{[t_1,t_2]}$ der Ereignisse im Zeitintervall $[t_1,t_2]$ mit $0 \le t_1 < t_2$ Poisson-verteilt mit

$$X_{[t_1,t_2]} = X_{t_2} - X_{t_1} \sim \text{Po}((t_2 - t_1)\lambda).$$

In die Verteilung geht also nur die Differenz $t_2 - t_1$ ein, nicht t_1 oder t_2 selbst.

Beispiel 7.18 (Serveranfragen)
Ein Server erhalte im Mittel 120 Anfragen pro Stunde. Wie hoch ist die Wahrscheinlichkeit, dass pro Minute mehr als 3 Anfragen eintreffen?

Ansatz für die Anzahl X_t der Anfragen im Zeitraum $[0,t]$ (in Minuten):

$$X_t \sim \text{Po}(t\lambda) \quad \text{mit} \quad \text{E}(X_{60}) = 60\lambda = 120 \quad \Rightarrow \quad \lambda = 2.$$

X_t: Anfragen im Zeitintervall $[0,t]$ min

$X_t \sim \text{Po}(t \cdot \lambda)$, $E(X_{60}) = 120$

$\Leftrightarrow 60 \cdot \lambda = 120$

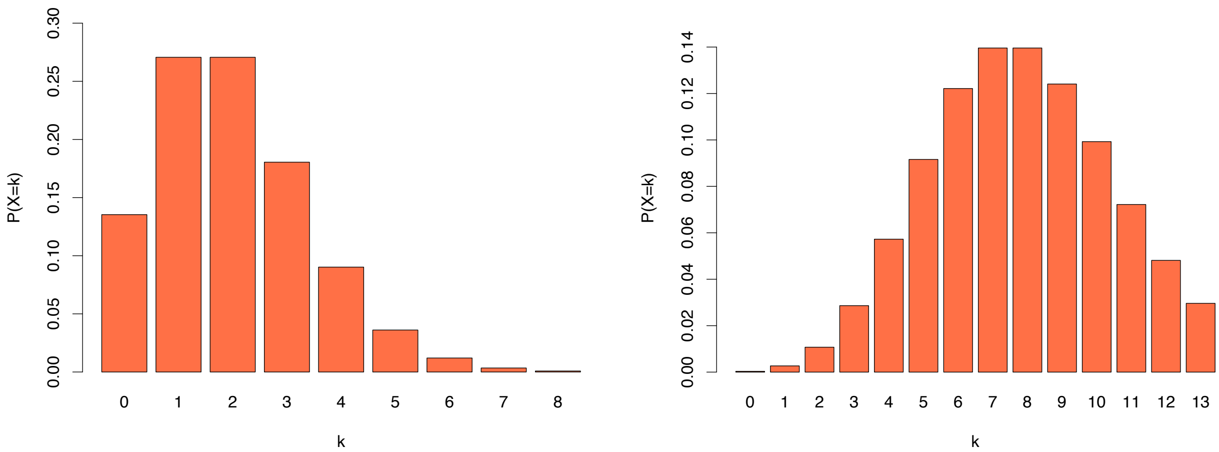
$\lambda = 2$

$$P(X_1 > 3) = 1 - P(X_1 \le 3) = 1 - \sum_{k=0}^{3} P(X_1 = k) = 1 - e^{-2} \sum_{k=0}^{3} \frac{2^k}{k!} = 0{,}143$$

Wegen $\mathrm{E}(X_1) = 1 \cdot \lambda$ können also 2 Anfragen pro Minute erwartet werden.

$$P(X_1 > 3) = 1 - P(X_1 \leq 3) = 1 - \sum_{k=0}^{3} P(X_1 = k) = 1 - e^{-2} \sum_{k=0}^{3} \frac{2^k}{k!} \approx 0.143\,.$$

Links: $X_1 \sim \mathrm{Po}(2)\,, \quad \mu = 2\,, \quad \sigma = \sqrt{2} \approx 1.41$
Rechts: $X_4 \sim \mathrm{Po}(8)\,, \quad \mu = 8\,, \quad \sigma = \sqrt{8} \approx 2.83$



Anfragen in 1 Minute (links) bzw. in 4 Minuten (rechts).

7.6 Übersicht

7.6.1 Eigenschaften diskreter Verteilungen

Name	Kurzschreibweise	Verteilung	Momente	
Bernoulli-Verteilung	$X \sim \mathrm{B}(1,p)$	$P(X=k) = \begin{cases} p & \text{falls} \quad k=1 \\ 1-p & \text{falls} \quad k=0 \end{cases}$	$\mathrm{E}(X) = p$ $\mathrm{Var}(X) = p\,(1-p)$	$n = 1$
Binomial-verteilung	$X \sim \mathrm{B}(n,p)$	$P(X=k) = \binom{n}{k} p^k\,(1-p)^{n-k}$	$\mathrm{E}(X) = n\,p$ $\mathrm{Var}(X) = n\,p\,(1-p)$	n fach durchführen
Hypergeom. Verteilung	$X \sim \mathrm{Hyp}(N,M,n)$	$P(X=k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$	$\mathrm{E}(X) = n\,\frac{M}{N}$ $\mathrm{Var}(X) = n\,\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$	
Poisson-Verteilung	$X \sim \mathrm{Po}(\lambda)$	$P(X=k) = \frac{\lambda^k}{k!}\,e^{-\lambda}$	$\mathrm{E}(X) = \lambda$ $\mathrm{Var}(X) = \lambda$	

7.6.2 Diskrete Verteilungen in R

Name	Kurzschreibweise	R-Funktionen

Bernoulli-Verteilung	$X \sim \mathrm{B}(1,p)$	`dbinom(x, 1, p)`
		`pbinom(q, 1, p)`
		`rbinom(size, 1, p)`

Binomial-verteilung	$X \sim \mathrm{B}(n,p)$	`dbinom(x, n, p)`
		`pbinom(q, n, p)`
		`rbinom(size, n, p)`

Hypergeom. Verteilung	$X \sim \mathrm{Hyp}(N,M,n)$	`dhyper(x, M, N-M, n)`
		`phyper(q, M, N-M, n)`
		`rhyper(size, M, N-M, n)`

Poisson-Verteilung	$X \sim \mathrm{Po}(\lambda)$	`dpois(x, lambda)`
		`ppois(q, lambda)`
		`rpois(size, lambda)`

* *Zu beachten:* Die R-Funktionen `dhyper()` etc. verwenden die Parameter M und $N-M$ statt (wie in der Kurzschreibweise) N und M.

- `d<name>(x, ...)`: Verteilung (p_i) zu einem Vektor (x_i) von Realisierungen $(x_i \in \mathbb{N}_0)$
- `p<name>(q, ...)`: Verteilungsfunktion $F(q_i)$ zu einem Vektor (q_i) von Quantilen $(q_i \in \mathbb{R})$
- `r<name>(size, ...)`: Vektor mit `size` Zufallszahlen gemäß zugehöriger Verteilung

```
barplot(dbinom(0:10, 10, 0.75)) # Binomialverteilung
random <- rpois(200, 5.0)       # 200 Po(5)-Zufallszahlen
table(random)                   # Abs. Haeufigkeit
```