Vorlesung 4a

Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Die Welt des p-Münzwurfs -

von Bernoulli zu Poisson



Jacob Bernoulli (1654-1705)

Sei
$$p \in (0,1)$$
, $q := 1 - p$.

Eine Bernoulli-Folge zum Parameter p

(man sagt manchmal auch: ein fortgesetzter p-Münzwurf(

ist eine Zufallsvariable $Z = (Z_1, Z_2, \ldots)$

mit Wertebereich $\{0,1\}^{\mathbb{N}}$,

deren Verteilung die folgende Eigenschaft hat:

Sei
$$p \in (0,1)$$
, $q := 1 - p$.

Eine Bernoulli-Folge zum Parameter p

(man sagt manchmal auch: ein fortgesetzter p-Münzwurf(

ist eine Zufallsvariable
$$Z = (Z_1, Z_2, ...)$$

mit Wertebereich $\{0, 1\}^{\mathbb{N}}$,

deren Verteilung die folgende Eigenschaft hat:

Für jedes $n \in \mathbb{N}$ und jede endliche 01-Folge (a_1, \ldots, a_n) mit k Einsen und n-k Nullen ist

$$P(Z_1 = a_1, ..., Z_n = a_n) = p^k q^{n-k}.$$

Zur Erinnerung:

Für jedes n ist dann die Anzahl der Einsen in (Z_1,\ldots,Z_n) (die "Anzahl der Erfolge in n Versuchen") binomial(n,p)-verteilt:

$$P(Z_1 + \dots + Z_n = k) = \binom{n}{k} p^k q^{n-k}.$$

$$T := \inf\{i : i \in \mathbb{N}, Z_i = 1\}$$

ist der Zeitpunkt des ersten Erfolges.

$$T := \inf\{i : i \in \mathbb{N}, Z_i = 1\}$$

ist der Zeitpunkt des ersten Erfolges.

Wie sieht die Verteilung von T aus?

$$P(T=n)=?$$

$$P(T = n) = ?$$

$${T = n} = {Z_1 = 0, ... Z_{n-1} = 0, Z_n = 1}$$

$$P(T=n)=?$$

$${T = n} = {Z_1 = 0, ... Z_{n-1} = 0, Z_n = 1}$$

Also:

$$P(T = n) = ?$$

$${T = n} = {Z_1 = 0, \dots Z_{n-1} = 0, Z_n = 1}$$

Also:

$$P(T = n) = P(Z_1 = 0, ... Z_{n-1} = 0, Z_n = 1)$$

$$P(T = n) = ?$$

$${T = n} = {Z_1 = 0, \dots Z_{n-1} = 0, Z_n = 1}$$

Also:

$$P(T = n) = P(Z_1 = 0, ... Z_{n-1} = 0, Z_n = 1)$$

$$= q^{n-1} p.$$

$$P(T=n)=?$$

$${T = n} = {Z_1 = 0, \dots Z_{n-1} = 0, Z_n = 1}$$

Also

$$P(T = n) = P(Z_1 = 0, ... Z_{n-1} = 0, Z_n = 1)$$

$$= q^{n-1} p.$$

Alternativ:

$${T > n} = {Z_1 = 0, \dots, Z_n = 0}$$

Alternativ:

$${T > n} = {Z_1 = 0, \dots, Z_n = 0}$$

Also

$$\mathbf{P}(T > n) = q^n.$$

$$P(T=n) = q^{n-1} p$$

$$\mathbf{P}(T > n) = q^n$$

$$\mathbf{P}(T=n) = q^{n-1} p$$

$$\mathbf{P}(T > n) = q^n$$

Das passt zusammen:

$$P(T = n) = P(T > n - 1) - P(T > n)$$

$$\mathbf{P}(T=n) = q^{n-1} p$$

$$P(T > n) = q^n$$

Das passt zusammen:

$$P(T = n) = P(T > n - 1) - P(T > n)$$

$$= q^{n-1} - q^{n}$$

$$= q^{n-1} (1 - q)$$

$$= q^{n-1} p.$$

Definition

Sei $p \in (0, 1)$. Eine Zufallsvariable T mit Zielbereich $\mathbb N$ heißt geometrisch verteilt mit Parameter p,

kurz Geom(p)-verteilt,

wenn

$$P(T > a) = q^a, \quad a = 0, 1, 2 \dots,$$

mit
$$q := 1 - p$$
.

$$\mathbf{E}[T] = ?$$

Anschaulich ist klar:

Beim gewöhnlichen Würfeln kommt im Mittel jedes 6-te Mal eine Sechs.

$$\mathbf{E}[T] = ?$$

Anschaulich ist klar:

Beim gewöhnlichen Würfeln kommt im Mittel jedes 6-te Mal eine Sechs.

Beim Münzwurf mit Erfolgswahrscheinlichkeit p kommt im Mittel jedes (1/p)-te Mal ein Erfolg.

$$\mathbf{E}[T] = ?$$

Anschaulich ist klar:

Beim gewöhnlichen Würfeln kommt im Mittel jedes 6-te Mal eine Sechs.

Beim Münzwurf mit Erfolgswahrscheinlichkeit p kommt im Mittel jedes (1/p)-te Mal ein Erfolg.

Also wird gelten:

$$\mathbf{E}[T] = \frac{1}{p}.$$

Das beweist man auch schnell mit dem folgenden

Lemma

(Buch S. 34)

Ist X eine Zufallsvariable mit Zielbereich \mathbb{N} oder \mathbb{N}_0 , dann ist

$$\mathbf{E}[X] = \sum_{i \ge 0} \mathbf{P}(X > i)$$

Beweis.

 $\rho(j)$ seien die Verteilungsgewichte von X.

$$E[X] = \sum_{j \ge 1} j \rho(j)$$

$$\sum_{i\geq 0} \mathbf{P}(X>i) = \sum_{i\geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$

Warum ist das gleich?

Beweis.

 $\rho(j)$ seien die Verteilungsgewichte von X.

$$\mathbf{E}[X] = \sum_{j \ge 1} j\rho(j) = \sum_{j \ge 1} \sum_{i=0}^{j-1} \rho(j)$$

$$\sum_{i\geq 0} \mathbf{P}(X>i) = \sum_{i\geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$

Warum ist das gleich?

Beweis.

 $\rho(j)$ seien die Verteilungsgewichte von X.

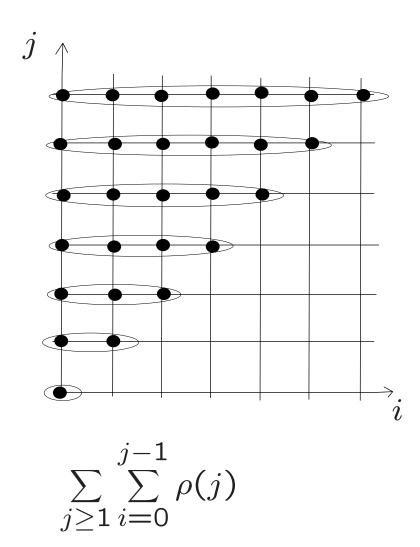
$$\mathbf{E}[X] = \sum_{j \ge 1} j\rho(j) = \sum_{j \ge 1} \sum_{i=0}^{j-1} \rho(j)$$

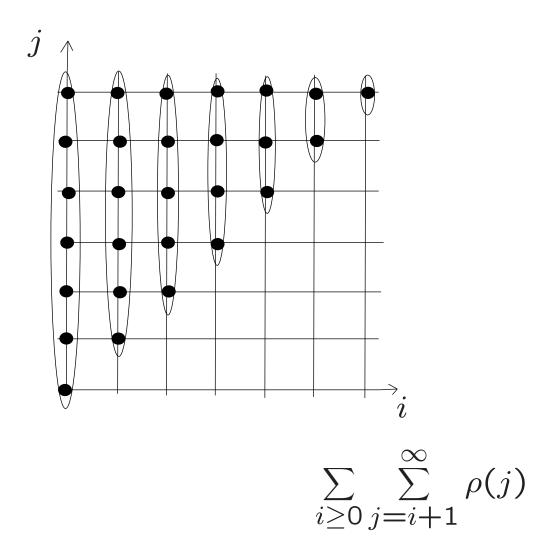
$$\sum_{i\geq 0} \mathbf{P}(X>i) = \sum_{i\geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$

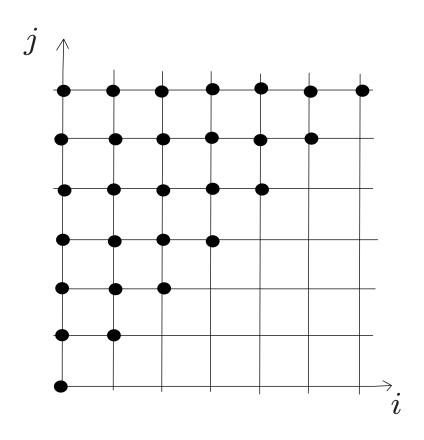
Warum ist das gleich?

Wie sieht man die Gleichheit

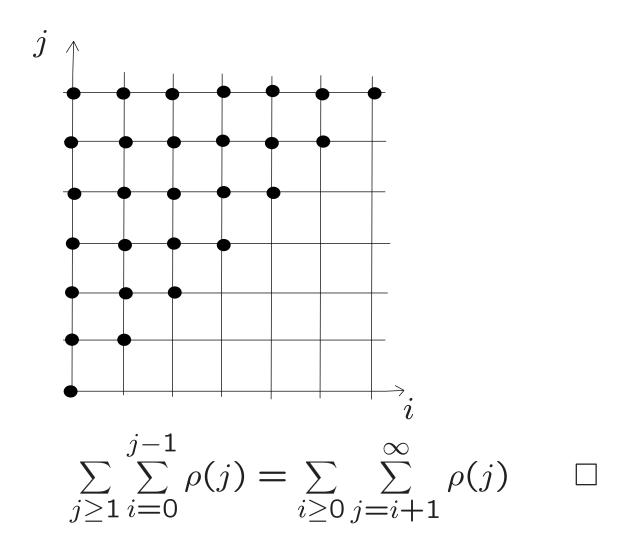
$$\sum_{j\geq 1} \sum_{i=0}^{j-1} \rho(j) = \sum_{i\geq 0} \sum_{j=i+1}^{\infty} \rho(j)$$
?







Es kommt nicht auf die Reihenfolge der Summation an



Fazit

Für eine Geom(p)-verteilte Zufallsvariable T ergibt sich:

$$E[T] = \sum_{i \ge 0} P(T > i) = \sum_{i \ge 0} q^i = \frac{1}{p}.$$

$$\mathbf{E}[T] = \frac{1}{p}$$

Nächster Akt:

Kleine Erfolgswahrscheinlichkeit:

Nächster Akt:

Kleine Erfolgswahrscheinlichkeit:

Wie lange dauert es bis zum ersten Erfolg?

Nächster Akt:

Kleine Erfolgswahrscheinlichkeit:

Wie lange dauert es bis zum ersten Erfolg?

Wieder sei

T

der zufällige Zeitpunkt des ersten Erfolgs in einem fortgesetzten p-Münzwurf.

$$p = \frac{1}{1000}$$

$$P(T > 2000) = ?$$

$$p = \frac{1}{1000}$$

$$P(T > 2000) = ?$$

$$q^{2000} = \left(1 - \frac{1}{1000}\right)^{2000} \approx$$

$$p = \frac{1}{1000}$$

$$P(T > 2000) = ?$$

$$q^{2000} = \left(1 - \frac{1}{1000}\right)^{2000} \approx e^{-2}$$

$$p = \frac{1}{1000}$$

$$P(T > 2000) \approx e^{-2}$$

$$q^{2000} = \left(1 - \frac{1}{1000}\right)^{2000} \approx e^{-2}$$

$$p = \frac{1}{1000}$$

$$P(T > 2000) \approx e^{-2}$$
.

$$p = \frac{1}{1000}$$

$$P(T > 2000) \approx e^{-2}$$
.

$$\mathbf{P}\left(\frac{T}{\mathbf{E}[T]} > 2\right) \approx e^{-2}.$$

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für
$$t \in \mathbb{R}_+$$
 ist

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für
$$t \in \mathbb{R}_+$$
 ist

$$\mathbf{P}\{\widetilde{T} > t\} = \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right)$$

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für
$$t \in \mathbb{R}_+$$
 ist

$$\mathbf{P}\{\widetilde{T} > t\} = \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right)$$
$$= \left(1 - p\right)^{\left\lfloor \frac{t}{p} \right\rfloor}$$

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für
$$t \in \mathbb{R}_+$$
 ist

$$\mathbf{P}\{\widetilde{T} > t\} = \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right)$$
$$= \left(1 - p\right)^{\left\lfloor \frac{t}{p} \right\rfloor}$$
$$= \left(1 - p\right)^{\frac{1}{p} p \left\lfloor \frac{t}{p} \right\rfloor}$$

$$\widetilde{T} := \frac{T}{\mathbf{E}[T]} = pT.$$

Für $t \in \mathbb{R}_+$ und kleines p ist

$$\mathbf{P}\{\widetilde{T} > t\} = \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right)$$
$$= \left(1 - p\right)^{\left\lfloor \frac{t}{p} \right\rfloor}$$
$$= \left(1 - p\right)^{\frac{1}{p} p \left\lfloor \frac{t}{p} \right\rfloor}$$
$$\approx (e^{-1})^t = e^{-t}.$$

Diese Tatsache formulieren wir als einen *Grenzwertsatz*:

(vgl. Buch S. 42)

Satz Sei X_1, X_2, \ldots eine Folge von

geometrisch verteilten Zufallsvariablen mit der Eigenschaft

$$\mathbf{E}[X_m] \underset{m \to \infty}{\longrightarrow} \infty.$$

Dann gilt für jedes $c \geq 0$:

Diese Tatsache formulieren wir als einen *Grenzwertsatz*:

(vgl. Buch S. 42)

Satz Sei X_1, X_2, \ldots eine Folge von

geometrisch verteilten Zufallsvariablen mit der Eigenschaft

$$\mathbf{E}[X_m] \underset{m \to \infty}{\longrightarrow} \infty.$$

Dann gilt für jedes c > 0:

$$\left| \mathbf{P} \left(\frac{X_m}{\mathbf{E}[X_m]} \ge c \right) \underset{m \to \infty}{\longrightarrow} e^{-c} \right|$$

Als nächstes betrachten wir wieder einen Münzwurf mit kleiner Erfolgswahrscheinlichkeit und fragen:

Wie viele Erfolge gibt es bei einer großen Zahl von Versuchen?

Als nächstes betrachten wir wieder einen Münzwurf mit kleiner Erfolgswahrscheinlichkeit und fragen:

Wie viele Erfolge gibt es bei einer großen Zahl von Versuchen?

p klein, n groß

$$X := Z_1 + Z_2 + \ldots + Z_n$$

$$P(X = k) \approx ?$$

$$p = \frac{1}{1000}, \quad n = 3000$$

$$p = \frac{1}{1000}, \quad n = 3000$$

$$P(X = 0) = q^n = \left(1 - \frac{1}{1000}\right)^{3000}$$

$$p = \frac{1}{1000}, \quad n = 3000$$

$$P(X = 0) = q^n = \left(1 - \frac{1}{1000}\right)^{3000}$$

 $\approx e^{-3}$

$$p = \frac{1}{1000}, \quad n = 3000$$

$$P(X = 0) = q^n = \left(1 - \frac{1}{1000}\right)^{3000}$$
$$\approx e^{-3}$$

$$P(X = 1) = npq^{n-1} \approx 3e^{-3}$$

$$p = \frac{1}{1000}, \quad n = 3000$$

$$P(X = 0) = q^n = \left(1 - \frac{1}{1000}\right)^{3000}$$

 $\approx e^{-3}$

$$P(X = 1) = npq^{n-1} \approx 3e^{-3}$$

$$P(X = 2) = \binom{n}{2} p^2 q^{n-2} \approx \frac{1}{2} (np)^2 q^n \approx \frac{1}{2} 3^2 e^{-3}$$

Clou:

p klein, n groß:

$$q^n = (1 - p)^n \approx e^{-np}$$

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{k!} n^k p^k q^n \approx \frac{1}{k!} (np)^k e^{-np}$$

Fazit

Sei p eine kleine positive Zahl, n eine große natürliche Zahl und X eine $\mathrm{Bin}(n,p)$ -verteilte Zufallsvariable.

Fazit

Sei p eine kleine positive Zahl, n eine große natürliche Zahl und X eine $\mathrm{Bin}(n,p)$ -verteilte Zufallsvariable.

Man kann dann die Verteilungsgewichte von X approximativ als Funktion von $\mathbf{E}[X] = n\,p\,$ ausdrücken. Rigoros fasst man diese Behauptung im folgenden Grenzwertsatz:

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n , $n = 1, 2, \ldots$,

eine Folge von $Bin(n, p_n)$ -verteilten Zufallsvariablen,

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n , $n = 1, 2, \ldots$,

eine Folge von $Bin(n, p_n)$ -verteilten Zufallsvariablen,

so dass für $n \to \infty$

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n , $n = 1, 2, \ldots$,

eine Folge von $Bin(n, p_n)$ -verteilten Zufallsvariablen,

so dass für $n \to \infty$

$$\mathbf{E}[X_n] o \lambda \;, \quad ext{d. h. } p_n \sim rac{\lambda}{n} \;.$$

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n , $n = 1, 2, \ldots$,

eine Folge von $Bin(n, p_n)$ -verteilten Zufallsvariablen,

so dass für $n \to \infty$

$$\mathbf{E}[X_n] o \lambda \;, \quad ext{d. h. } p_n \sim rac{\lambda}{n} \;.$$

Dann gilt für jedes $k = 0, 1, 2, \dots$

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei X_n , n = 1, 2, ...,

eine Folge von $Bin(n, p_n)$ -verteilten Zufallsvariablen,

so dass für $n \to \infty$

$$\mathbf{E}[X_n] o \lambda \;, \quad \mathsf{d.\,h.} \;\; p_n \sim rac{\lambda}{n} \;.$$

Dann gilt für jedes $k = 0, 1, 2, \dots$

$$\mathbf{P}(X_n = k) \to \frac{\lambda^k}{k!} e^{-\lambda} .$$



Siméon Denis Poisson (1781-1840)

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\binom{n}{k}p_n^k(1-p_n)^{n-k} =$$

$$\frac{1}{k!} \cdot \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot (np_n)^k \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k}.$$

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\frac{1}{k!} \cdot \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\longrightarrow 1} \cdot (np_n)^k \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k}$$

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\frac{1}{k!} \cdot \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\longrightarrow 1} \cdot \underbrace{(np_n)^k}_{\longrightarrow \lambda^k} \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k}$$

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\frac{1}{k!} \cdot \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\rightarrow 1} \underbrace{\frac{(np_n)^k}{h^k}}_{\rightarrow k} \underbrace{\left(1 - \frac{np_n}{n}\right)^n}_{\rightarrow e^{-\lambda}} (1 - p_n)^{-k}$$

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\frac{1}{k!} \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\rightarrow 1} \underbrace{\frac{(np_n)^k}{h^k}}_{\rightarrow k} \underbrace{\left(1 - \frac{np_n}{n}\right)^n}_{\rightarrow e^{-\lambda}} \underbrace{\left(1 - \frac{p_n}{n}\right)^{-k}}_{\rightarrow 1}$$

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} =$$

$$\frac{1}{k!} \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\longrightarrow 1} \underbrace{(np_n)^k}_{\longrightarrow \lambda^k} \underbrace{\left(1 - \frac{np_n}{n}\right)^n}_{\longrightarrow e^{-\lambda}} \underbrace{(1 - p_n)^{-k}}_{\longrightarrow 1}$$

$$\rightarrow \frac{1}{k!} \lambda^k e^{-\lambda}$$
. \square

(Buch S. 29)

Sei $\lambda \in \mathbb{R}_+$.

Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

(Buch S. 29)

Sei $\lambda \in \mathbb{R}_+$.

Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

Poissonverteilt mit Parameter λ ,

(Buch S. 29)

Sei $\lambda \in \mathbb{R}_+$.

Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

Poissonverteilt mit Parameter λ ,

kurz Pois(λ)-verteilt,

(Buch S. 29)

Sei $\lambda \in \mathbb{R}_+$.

Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

Poissonverteilt mit Parameter λ ,

kurz Pois(λ)-verteilt,

wenn

(Buch S. 29)

Sei
$$\lambda \in \mathbb{R}_+$$
.

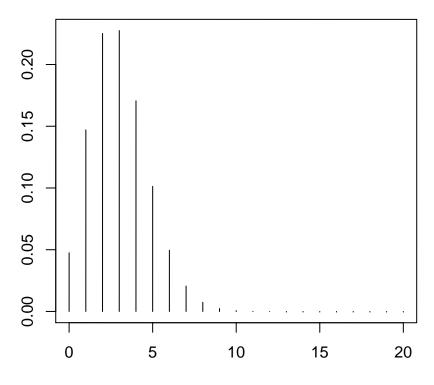
Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

Poissonverteilt mit Parameter λ ,

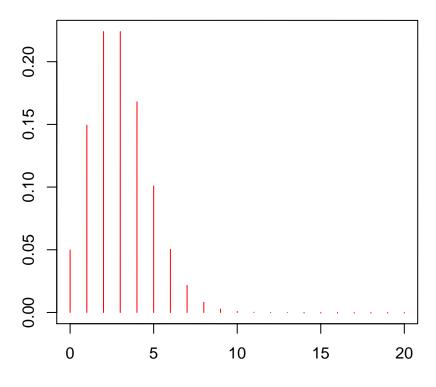
kurz Pois(λ)-verteilt,

wenn

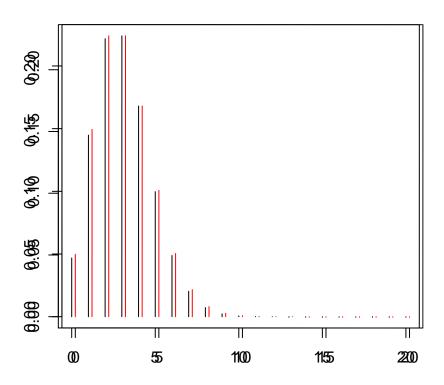
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$



Binomialgewichte zu $n=100~{\rm und}~p=0.03$



Poissongewichte zum Parameter λ =3



$\label{eq:definition} \mbox{Der Erwartungswert}$ einer $\mbox{Pois}(\lambda)\mbox{-verteilten Zufallsvariablen }X$ ist

Der Erwartungswert einer $\operatorname{Pois}(\lambda)$ -verteilten Zufallsvariablen X ist $\mathbf{E}[X] = \lambda.$

 $\label{eq:definition} \mbox{Der Erwartungswert}$ einer $\mbox{Pois}(\lambda)\mbox{-verteilten Zufallsvariablen }X$ ist

$$\mathbf{E}[X] = \lambda.$$

Der Erwartungswert einer $Pois(\lambda)$ -verteilten Zufallsvariablen X ist

$$\mathbf{E}[X] = \lambda.$$

$$\mathbf{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda}$$

Der Erwartungswert einer $Pois(\lambda)$ -verteilten Zufallsvariablen X ist

$$E[X] = \lambda.$$

$$\mathbf{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda \cdot 1 \qquad \Box$$

Zusammenfassung:

1. Im p-Münzwurf ist die Wartezeit auf den ersten Erfolg

Geom(p)-verteilt: $P(T > n) = q^n$.

Zusammenfassung:

1. Im p-Münzwurf ist die Wartezeit auf den ersten Erfolg

Geom(p)-verteilt:
$$P(T > n) = q^n$$
.

2. Für kleine
$$p$$
 gilt: $\mathbf{P}\left(\frac{T}{\mathbf{E}[T]} > t\right) \approx e^{-t}$

Zusammenfassung:

1. Im p-Münzwurf ist die Wartezeit auf den ersten Erfolg Geom(p)-verteilt: $\mathbf{P}(T>n)=q^n$.

2. Für kleine
$$p$$
 gilt: $\mathbf{P}\left(\frac{T}{\mathbf{E}[T]} > t\right) \approx e^{-t}$

3. Für kleine p und große n ist die Anzahl der Erfolge in n Versuchen approximativ Pois(np)-verteilt.

Für eine $Pois(\lambda)$ -verteilte Zufallsvariable gilt:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$