

Vorlesung 5a

Versuche, Erfolge, Wartezeiten:

Die Welt des p -Münzwurfs -

von Bernoulli zu Poisson

0. Fortgesetzter p -Münzwurf

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

Eine *Bernoulli-Folge zum Parameter p*

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

ist eine zufällige 01-Folge (Z_1, Z_2, \dots) , deren Verteilung die folgende Eigenschaft hat:

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

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

(d.h. für jedes n ist (Z_1, \dots, Z_n) ein n -facher p -Münzwurf)

Zur Erinnerung:

Für jedes n ist dann

die Anzahl der Einsen in (Z_1, \dots, Z_n)
(die “Anzahl der Erfolge in n Versuchen”)

binomial(n, p)-verteilt:

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

1. Der Zeitpunkt des ersten Erfolgs
und
die geometrische Verteilung

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

ist der Zeitpunkt des ersten Erfolges.

Wie sieht die Verteilung von T aus?

$$\mathbf{P}(T = n) = ?$$

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

Also

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

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

Alternativ:

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

Also

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

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

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

Das passt auch zusammen:

$$\begin{aligned} \mathbf{P}(T = n) &= \mathbf{P}(T > n - 1) - \mathbf{P}(T > n) \\ &= q^{n-1} - q^n \\ &= q^{n-1} (1 - q) \\ &= q^{n-1} p. \end{aligned}$$

Definition

Sei $p \in (0, 1)$. Eine Zufallsvariable T mit Zielbereich \mathbb{N} heißt

geometrisch verteilt mit Parameter p ,

kurz **Geom(p)-verteilt**,

wenn

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

mit $q := 1 - p$.

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

Anschaulich ist klar:

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

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

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

Beweis.

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

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

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

Warum sind die beiden **rechten Seiten** gleich?

Beweis.

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

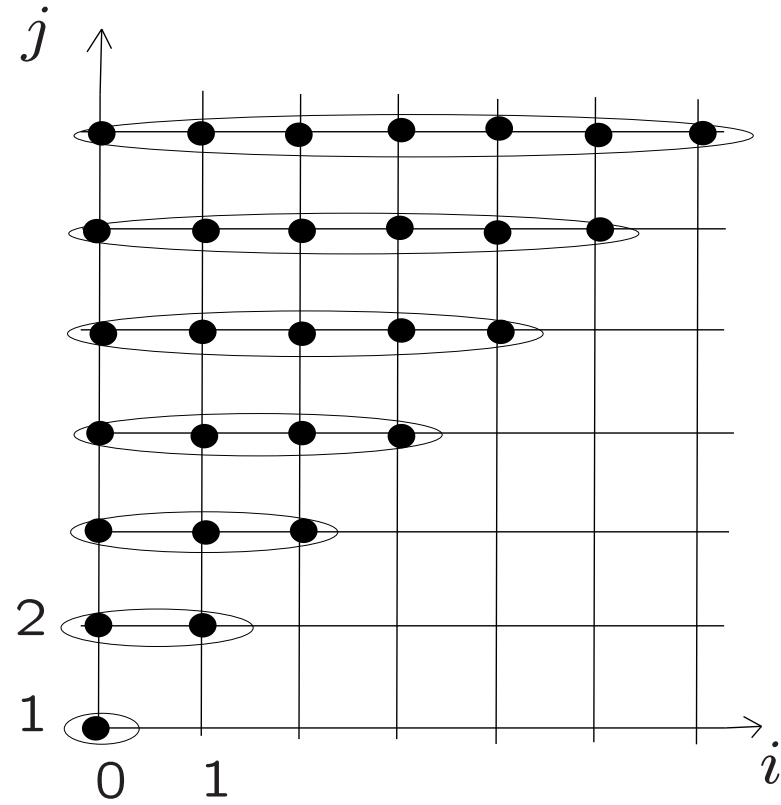
$$\mathbf{E}[X] = \sum_{j \geq 1} j \rho(j) = \sum_{j \geq 1} \sum_{i=0}^{j-1} \rho(j)$$

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

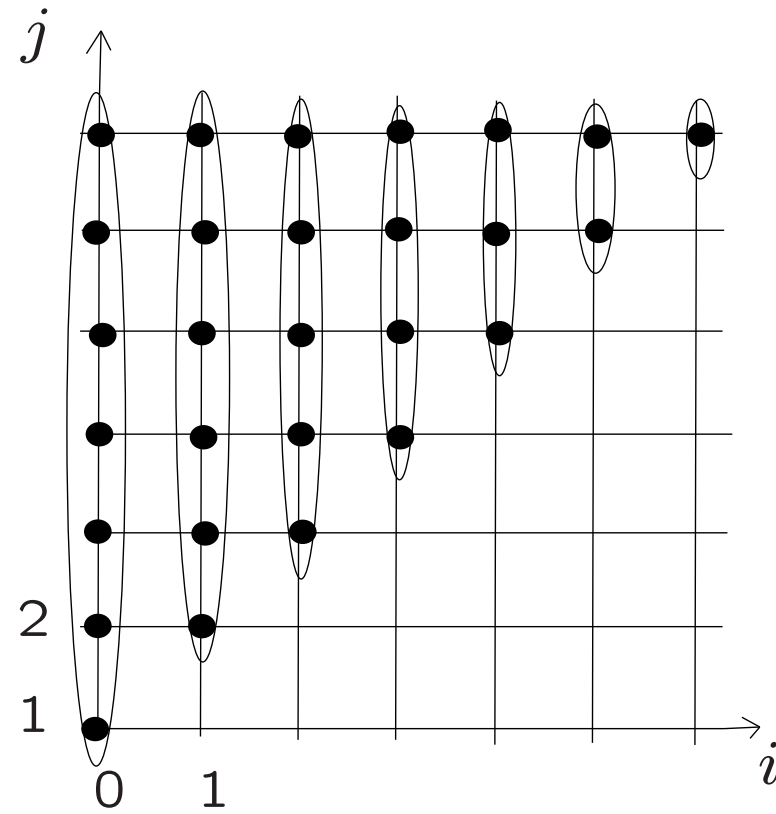
Warum sind die beiden **rechten Seiten** gleich?

Wie sieht man die Gleichheit

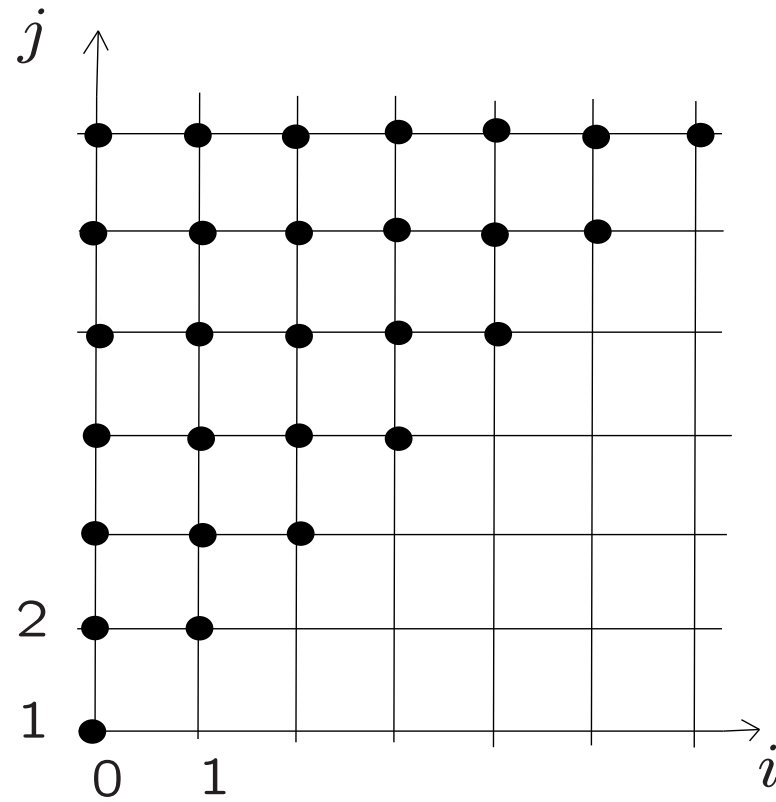
$$\sum_{j \geq 1} \sum_{i=0}^{j-1} \rho(j) = \sum_{i \geq 0} \sum_{j=i+1}^{\infty} \rho(j) \quad ?$$



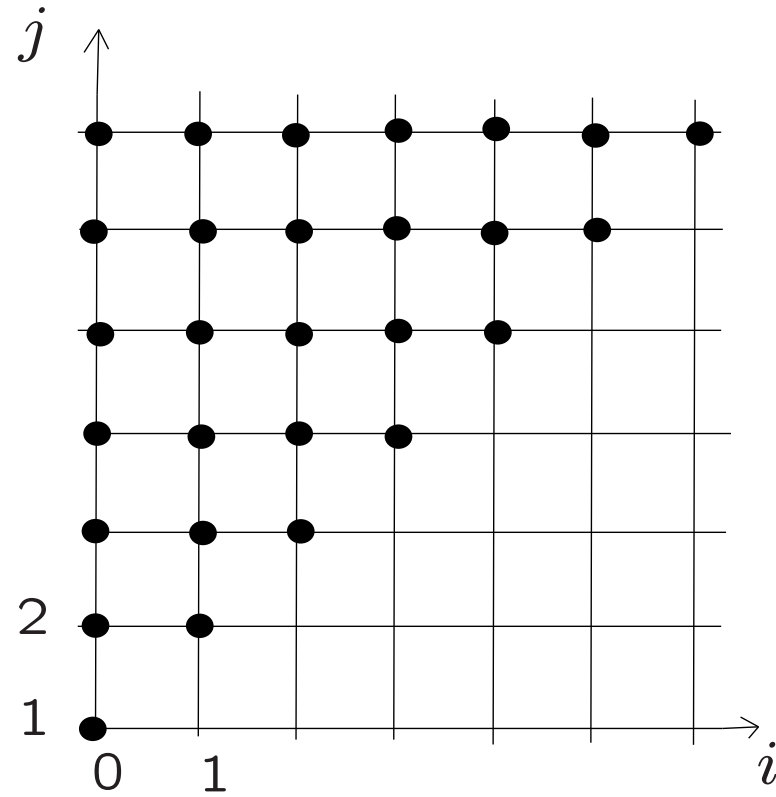
$$\sum_{j \geq 1} \sum_{i=0}^{j-1} \rho(j)$$



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



Es kommt nicht auf die Reihenfolge
der Summation an



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

Fazit

Für eine $\text{Geom}(p)$ -verteilte Zufallsvariable T ergibt sich:

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

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

2. Die Exponentialapproximation.

Oder:

Münzwurf

mit kleiner Erfolgswahrscheinlichkeit:

Wie lange dauert es bis zum ersten Erfolg?

Wieder sei

T

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

Beispiel:

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

$$\mathbf{P}(T > 2000) = ?$$

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

Beispiel:

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

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

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

Beispiel:

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

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

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

Betrachten wir T auf der Skala seines Erwartungswertes:

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

Für $t \in \mathbb{R}_+$ ist

$$\begin{aligned} \mathbf{P}\{\tilde{T} > t\} &= \mathbf{P}\left(T > \frac{t}{p}\right) = \mathbf{P}\left(T > \left\lfloor \frac{t}{p} \right\rfloor\right) \\ &= (1 - p)^{\left\lfloor \frac{t}{p} \right\rfloor} \\ &= (1 - p)^{\frac{1}{p}} p^{\left\lfloor \frac{t}{p} \right\rfloor}. \end{aligned}$$

Für $p \rightarrow 0$ konvergiert dies gegen

$$(e^{-1})^t = e^{-t}.$$

Diese Tatsache formulieren wir als einen *Grenzwertsatz*:

(vgl. Buch S. 42)

Satz Sei X_1, X_2, \dots eine Folge von geometrisch verteilten Zufallsvariablen mit der Eigenschaft

$$\mathbf{E}[X_m] \xrightarrow{m \rightarrow \infty} \infty.$$

Dann gilt für jedes $t \geq 0$:

$$\mathbf{P} \left(\frac{X_m}{\mathbf{E}[X_m]} > t \right) \xrightarrow{m \rightarrow \infty} e^{-t}$$

3. Die Poissonapproximation.

Oder

Münzwurf mit kleiner Erfolgswahrscheinlichkeit:

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

p klein, n groß

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

$$\mathbf{P}(X = k) \approx ?$$

Beispiel:

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

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

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

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

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 $\text{Bin}(n, p)$ -verteilte Zufallsvariable.

Man kann dann die Verteilungsgewichte von X
approximativ als Funktion von $\mathbb{E}[X] = np$ ausdrücken.

Rigoros fasst man diese Behauptung im folgenden

Grenzwertsatz:

Satz (Poissons Gesetz der seltenen Ereignisse)

(vgl. Buch S. 30)

Sei $\lambda > 0$ und sei $X_n, n = 1, 2, \dots$,
eine Folge von $\text{Bin}(n, p_n)$ -verteilten Zufallsvariablen,

so dass für $n \rightarrow \infty$

$$\mathbf{E}[X_n] \rightarrow \lambda, \quad \text{d. h. } p_n \sim \frac{\lambda}{n}.$$

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

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

Beweis:

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

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

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

Definition (Poissonverteilung)

(Buch S. 29)

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

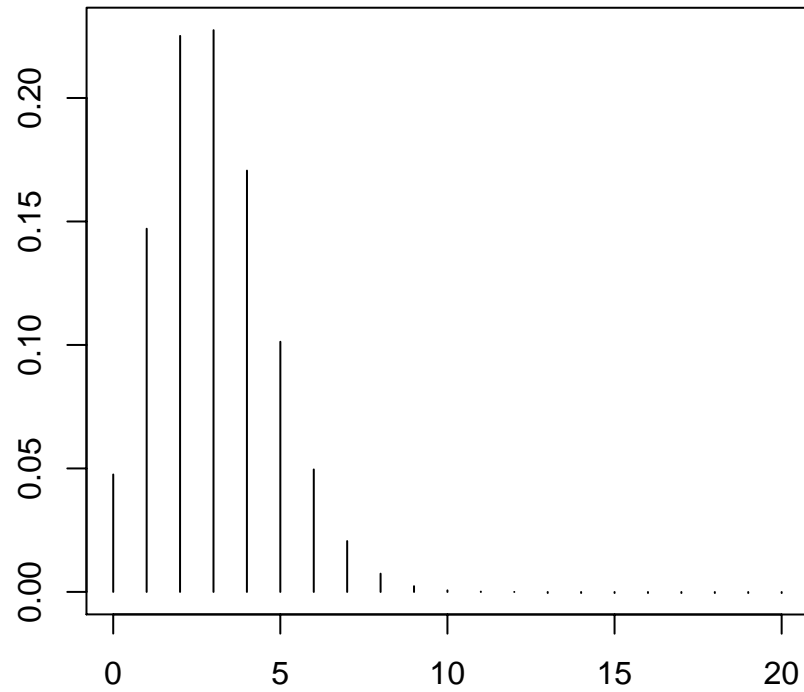
Eine Zufallsvariable X mit Zielbereich \mathbb{N}_0 heißt

Poissonverteilt mit Parameter λ ,

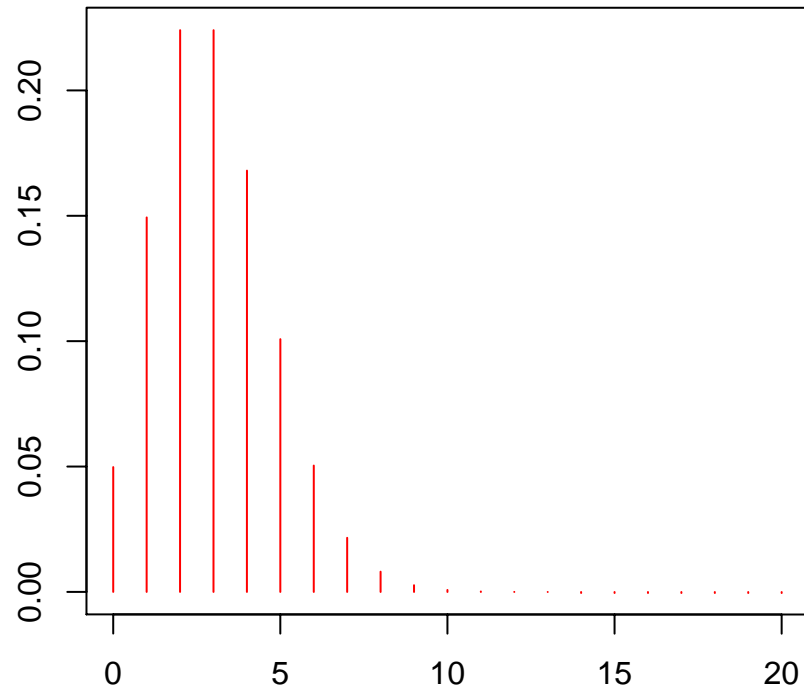
kurz $\text{Pois}(\lambda)$ -verteilt,

wenn

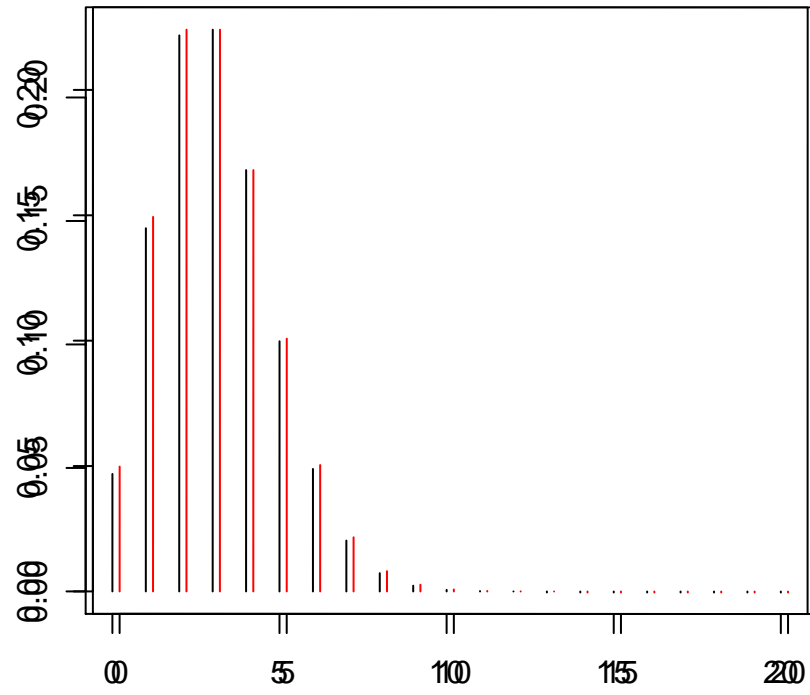
$$\mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$



Binomialgewichte zu $n = 100$ und $p = 0.03$



Poissongewichte zum Parameter $\lambda = 3$



Satz.

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

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

Beweis:

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

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

Zusammenfassung:

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

$$\text{Geom}(p)\text{-verteilt: } \mathbf{P}(T > n) = q^n.$$

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 $\text{Pois}(np)$ -verteilt.

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

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