Vorlesung 4a

Versuche, Erfolge, Wartezeiten:

Die Welt des p-Münzwurfs -

von Bernoulli zu Poisson

Jacob Bernoulli (1654-1705)

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

$(Z_1, Z_2, \ldots)$ ein *fortgesetzter* p*-Münzwurf*

(eine *Bernoulli-Folge* zum Parameter p) :

Für jede endliche 01-Folge $(a_1, \ldots, a_n)$

mit k Einsen und $n - k$ Nullen ist

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

Zur Erinnerung:

Für jedes n ist

die Anzahl der Einsen in $(Z_1, \dots, Z_n)$

(die “Anzahl der Erfolge in n Versuchen”)

binomial(n, p)-verteilt:

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

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

ist der Zeitpunkt des ersten Erfolges.

Wie sieht die Verteilung von T aus?

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

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

Also

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

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

Alternativ:

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

Also

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

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

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

Das passt zusammen:

$$\mathbf{P}(T = n) = \mathbf{P}(T > n-1) - \mathbf{P}(T > n)$$

$$\begin{aligned} &= q^{n-1} - q^{n} \\ &= q^{n-1}\,(1-q) \\ &= q^{n-1}\, p. \end{aligned}$$

Definition

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

kurz Geom(p)-verteilt,

wenn

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

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.

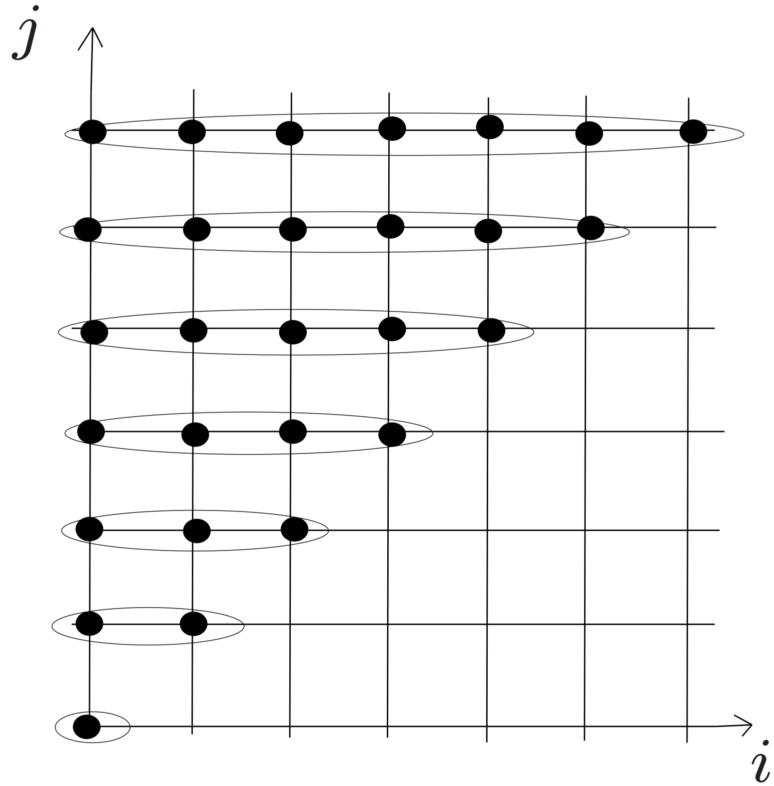
$$\mathbf{E}[X] = \sum_{j\geq 1} j\rho(j) = \sum_{j\geq 1}\sum_{i=0}^{j-1} \rho(j)$$

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

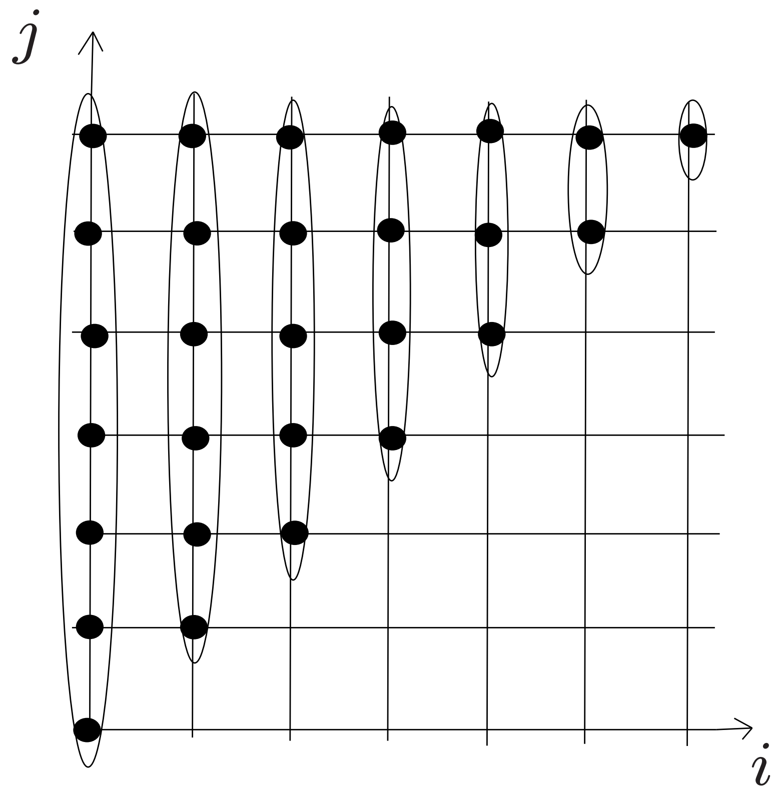
Warum ist das gleich?

Wie sieht man die Gleichheit

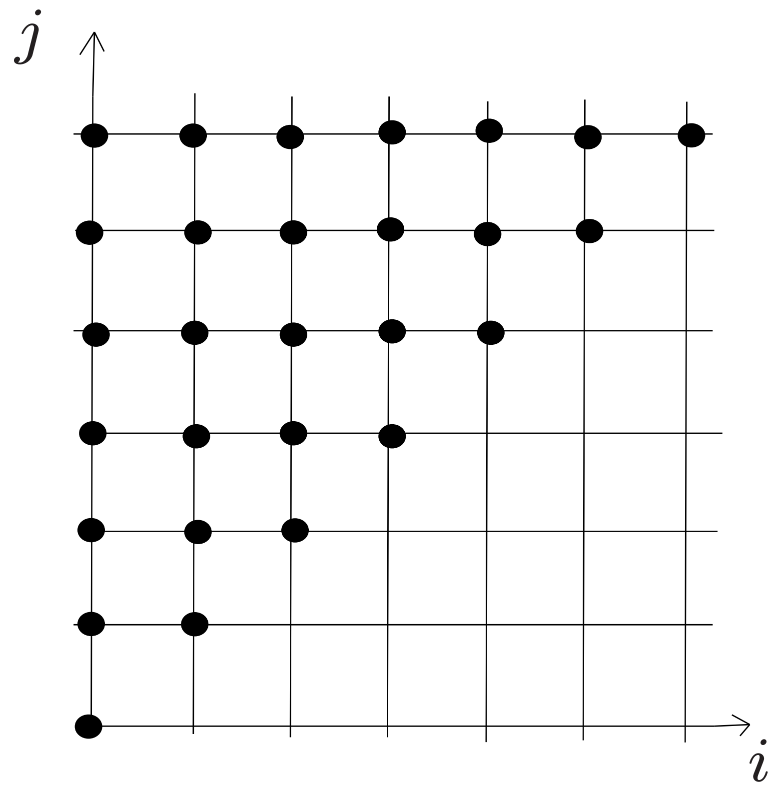
$$\sum_{j\geq 1}\sum_{i=0}^{j-1}\rho(j) = \sum_{i\geq 0}\sum_{j=i+1}^{\infty}\rho(j) \qquad ?$$



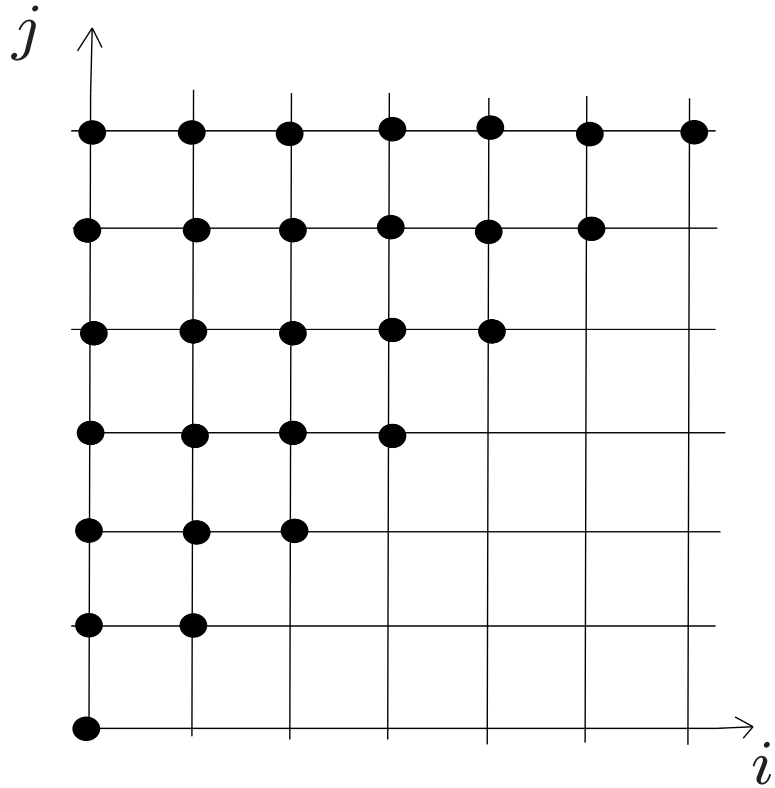
$$\sum_{j \geq 1} \sum_{i=0}^{j-1} \rho(j)$$



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



Es kommt nicht auf die Reihenfolge
der Summation an



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

Fazit

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

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

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

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.

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*:

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

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 $c \geq 0$:

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

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, $\quad n$ groß

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

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

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 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:

Satz (Poissons Gesetz der seltenen Ereignisse)

(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 $n \to \infty$

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

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

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

Siméon Denis Poisson (1781-1840)

Beweis:

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

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

$$\to \frac{1}{k!}\lambda^k e^{-\lambda}\,. \qquad \square$$

Definition (Poissonverteilung)

(Buch S. 29)

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

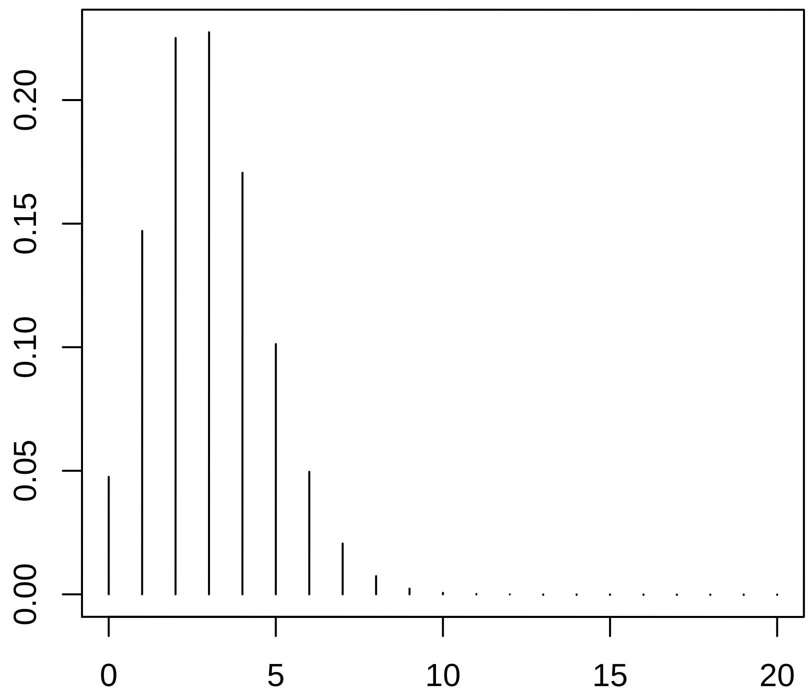
Eine Zufallsvariable X mit Zielbereich $\mathbb{N}_0$ heißt

Poissonverteilt mit Parameter λ,

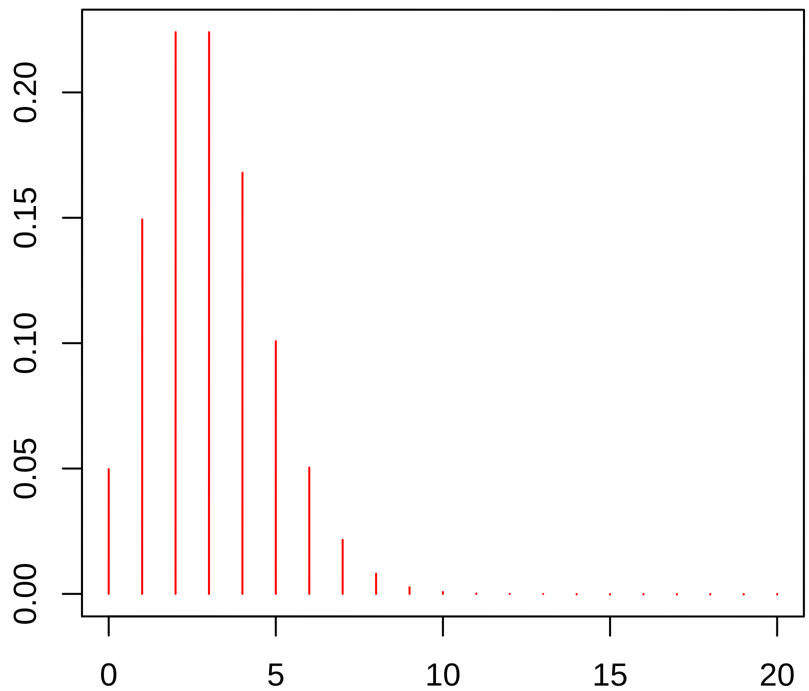
kurz Pois(λ)-verteilt,

wenn

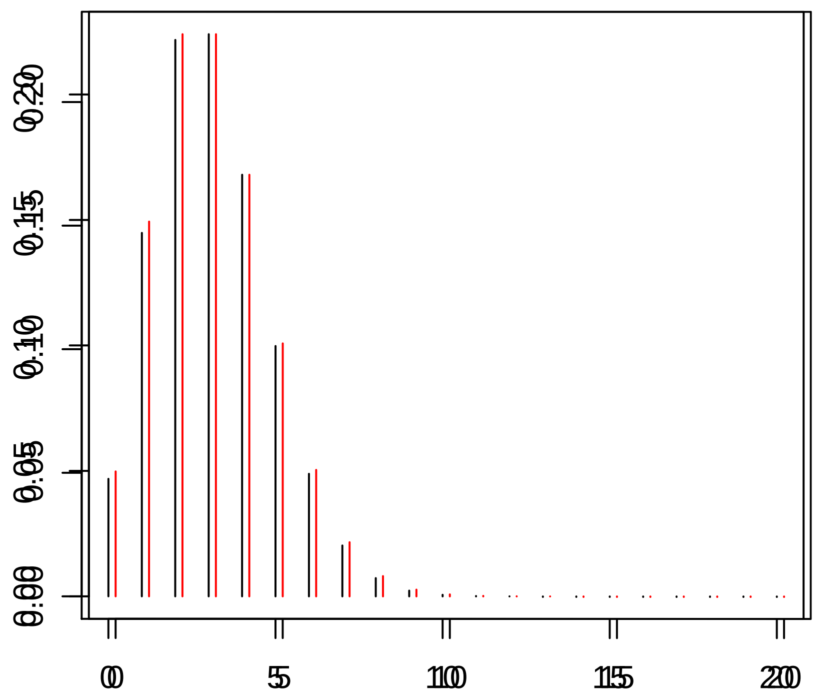
$$\boxed{\mathbf{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots .}$$



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



Poissongewichte zum Parameter λ =3



0.00
0.05
0.10
0.15
0.20
0
5
10
15
20

Satz.

Der Erwartungswert

einer Pois(λ)-verteilten Zufallsvariablen X ist

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

Beweis:

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

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

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(λ)-verteilte Zufallsvariable gilt:

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