Statistik für Informatiker, SS 2017

Zur Poissonapproximation der Binomialverteilung

Matthias Birkner

http://www.staff.uni-mainz.de/birkner/StatInfo17/

17.5.2017

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Erinnerung

Beispiel 1.21 (Poissonverteilung)

$\lambda \in (0, \infty)$,

$$\mathrm{Poi}_\lambda\big(\{k\}\big) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k \in \mathbb{N}_0$$

Proposition 1.22 (Poissonapproximation der Binomialverteilung)

Seien $p_n \in [0,1]$ *mit* $p_n \to 0$ *und* $np_n \to \lambda \in (0,\infty)$ *für* $n \to \infty$*, so gilt für jedes* $k \in \mathbb{N}_0$

$$\mathrm{Bin}_{n,p_n}\big(\{k\}\big) \underset{n\to\infty}{\longrightarrow} \mathrm{Poi}_\lambda\big(\{k\}\big).$$

Proposition 1.22 (Poissonapproximation der Binomialverteilung)

Seien $p_n \in [0,1]$ mit $p_n \to 0$ und $np_n \to \lambda \in (0,\infty)$ für $n \to \infty$, so gilt für jedes $k \in \mathbb{N}_0$

$$\mathrm{Bin}_{n,p_n}\big(\{k\}\big) \xrightarrow[n\to\infty]{} \mathrm{Poi}_\lambda\big(\{k\}\big).$$

Beweis. Es ist

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

Proposition 1.22 (Poissonapproximation der Binomialverteilung)

Seien $p_n \in [0,1]$ mit $p_n \to 0$ und $np_n \to \lambda \in (0,\infty)$ für $n \to \infty$, so gilt für jedes $k \in \mathbb{N}_0$

$$\mathrm{Bin}_{n,p_n}\big(\{k\}\big) \xrightarrow[n\to\infty]{} \mathrm{Poi}_\lambda\big(\{k\}\big).$$

Beweis. Es ist

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \underbrace{\frac{n(n-1)\cdots(n-k+1)}{k!\,n^k}}_{\to 1/k!} \big(\underbrace{np_n}_{\to\lambda}\big)^k \underbrace{\Big(1-\frac{np_n}{n}\Big)^n}_{\to e^{-\lambda}} (1-p_n)^{-k}$$

Proposition 1.22 (Poissonapproximation der Binomialverteilung)

Seien $p_n \in [0,1]$ mit $p_n \to 0$ und $np_n \to \lambda \in (0,\infty)$ für $n \to \infty$, so gilt für jedes $k \in \mathbb{N}_0$

$$\mathrm{Bin}_{n,p_n}\big(\{k\}\big) \underset{n\to\infty}{\longrightarrow} \mathrm{Poi}_\lambda\big(\{k\}\big).$$

Beweis. Es ist

$$\binom{n}{k} p_n^k (1-p_n)^{n-k}$$
$$= \underbrace{\frac{n(n-1)\cdots(n-k+1)}{k!\, n^k}}_{\to 1/k!} \big(\underbrace{np_n}_{\to\lambda}\big)^k \underbrace{\left(1-\frac{np_n}{n}\right)^n}_{\to e^{-\lambda}} (1-p_n)^{-k}$$
$$\longrightarrow e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{für } n\to\infty.$$

Prop. 1.22:

$$np_n \to \lambda \in (0,\infty), \quad \text{so gilt} \quad \mathrm{Bin}_{n,p_n}(k) \underset{n\to\infty}{\longrightarrow} \mathrm{Poi}_\lambda(k)$$

Prop. 1.22 motiviert, warum die Poissonverteilung oft in Anwendungssituationen vorkommt, in denen man viele unabhängige Ereignisse betrachtet, von denen jedes nur mit einer sehr kleinen W'keit eintritt.

Man denke etwa an Schädensfälle bei Versicherungen, Zerfallsereignisse in einer Probe radioaktiven Materials oder an genetische Mutationen.

Prop. 1.22:

$$np_n \to \lambda \in (0,\infty), \quad \text{so gilt} \quad \mathrm{Bin}_{n,p_n}(k) \xrightarrow[n\to\infty]{} \mathrm{Poi}_\lambda(k)$$

Prop. 1.22 motiviert, warum die Poissonverteilung oft in Anwendungssituationen vorkommt, in denen man viele unabhängige Ereignisse betrachtet, von denen jedes nur mit einer sehr kleinen W'keit eintritt.

Man denke etwa an Schädensfälle bei Versicherungen, Zerfallsereignisse in einer Probe radioaktiven Materials oder an genetische Mutationen.

Beispiel 1.23

L. von von Bortkewitschberichtete in seinem Buch *Das Gesetz der kleinen Zahlen*, Teubner, 1898 verschiedene Datensätze, die gut zur Poissonverteilung passen.

Beispiel 1.23

Speziell in § 12, 4. („Die durch Schlag eines Pferdes im preußischen Heere getöteten") werden für 20 Jahre (1875–1894) und 10 Armeekops der preußischen Kavallerie, also insgesamt $20 \cdot 10 = 200$ „Korpsjahre" berichtet, in wievielen davon sich x Todesfälle durch Schlag eines Pferds ereigneten (Tabelle b) auf S. 25):

Ergebnis x	Anz. „Korpsjahre"
0	109
1	65
2	22
3	3
4	1
≥ 5	0

Beispiel 1.23, Forts.

Angenommen, die Anzahl durch Schlag eines Pferdes während eines Jahres in einem Korps getöteter Soldaten wäre Poi_λ-verteilt mit $\lambda = 0{,}61$, so würden wir das Resultat x je $200 \times \mathrm{Poi}_{0,61}(x)$-mal erwarten:

Ergebnis x	Anz. „Korpsjahre“	$200 \times \mathrm{Poi}_{0,61}(x)$
0	109	108,67
1	65	66,29
2	22	20,22
3	3	4,11
4	1	0,63
≥ 5	0	0,08

Beispiel 1.23, Forts.

Angenommen, die Anzahl durch Schlag eines Pferdes während eines Jahres in einem Korps getöteter Soldaten wäre Poi_λ-verteilt mit $\lambda = 0{,}61$, so würden wir das Resultat x je $200 \times \mathrm{Poi}_{0,61}(x)$-mal erwarten:

Ergebnis x	Anz. „Korpsjahre“	$200 \times \mathrm{Poi}_{0,61}(x)$
0	109	108,67
1	65	66,29
2	22	20,22
3	3	4,11
4	1	0,63
≥ 5	0	0,08

Von Bortkewitsch, a.a.O., S. 25 schreibt: „Die Kongruenz der Theorie mit der Erfahrung lässt [...], wie man sieht, nichts zu wünschen übrig.“

Beispiel 1.23, Forts. 2

Übrigens: Wie ist von Bortkewitsch auf $\lambda = 0{,}61$ gekommen?

Beispiel 1.23, Forts. 2

Übrigens: Wie ist von Bortkewitsch auf $\lambda = 0{,}61$ gekommen?

Die beobachtete „mittlere Anzahl Todesfälle pro Korpsjahr“ in den Daten ist

$$\widehat{\lambda} = \frac{109}{200} \cdot 0 + \frac{65}{200} \cdot 1 + \frac{22}{200} \cdot 2 + \frac{3}{200} \cdot 3 + \frac{1}{200} \cdot 4 + 0 = 0{,}61$$

Beispiel 1.23, Forts. 2

Übrigens: Wie ist von Bortkewitsch auf $\lambda = 0{,}61$ gekommen?

Die beobachtete „mittlere Anzahl Todesfälle pro Korpsjahr“ in den Daten ist

$$\widehat{\lambda} = \frac{109}{200} \cdot 0 + \frac{65}{200} \cdot 1 + \frac{22}{200} \cdot 2 + \frac{3}{200} \cdot 3 + \frac{1}{200} \cdot 4 + 0 = 0{,}61$$

und es ist auch

$$\sum_{x=0}^{\infty} x \mathrm{Poi}_\lambda(x) = \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} = \lambda$$

(„der Erwartungswert von Poi_λ ist λ“) und somit ist obiges der naheliegende „Momentenschätzer“ – wir werden darauf zurückkommen.