

66.9 Schwaches Gesetz der großen Zahlen

Mit Hilfe der Tschebyschew-Ungleichung kann man das schwache Gesetz der großen Zahlen beweisen:

Es werde ein Versuch n Mal wiederholt, bei dem das Ereignis A mit Wahrscheinlichkeit p eintritt. Dann strebt die Wahrscheinlichkeit, dass sich die relative Häufigkeit $h_n(A)$ um weniger als eine beliebig kleine Zahl $\varepsilon > 0$ von p unterscheidet gegen 1 für $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} P(|h_n(A) - p| < \varepsilon) = 1$$

§67: WICHTIGE DISKRETE VERTEILUNGEN

67.1 Motivation:

Einige diskrete Verteilungen treten sehr häufig auf und tragen einen eigenen Namen. Wir wollen 4 dieser Verteilungen genauer betrachten: die Gleichverteilung, die Binomialverteilung, die Poisson-Verteilung und die geometrische Verteilung.

67.2 Die Gleichverteilung

Sei $\Omega = \{\omega_1, \dots, \omega_n\}$ ein Stichprobenraum mit $P(\omega_k) = \frac{1}{n}$ für alle $k = 1, \dots, n$. Ferner sei U_n eine Zufallsvariable mit $U_n(\omega_k) = k$. Dann ist $P_{U_n}(k) = \frac{1}{n}$. Eine solche Verteilung heißt (diskrete) Gleichverteilung auf der Menge $\{1, \dots, n\}$.

Es gilt:

$$\mu = E(U_n) = \sum_{k=1}^n k P_{U_n}(k) = \frac{1}{n} \cdot \sum_{k=1}^n k$$

$$\sigma^2 = V(U_n) = E(U_n^2) - \mu^2 = \sum_{k=1}^n k^2 \cdot P_{U_n}(k) - \mu^2 = \frac{1}{n} \sum_{k=1}^n k^2 - \frac{1}{n^2} \left(\sum_{k=1}^n k \right)^2$$

67.3 Beispiel:

Beim Würfeln liegt eine diskrete Gleichverteilung auf der Menge $\{1, \dots, 6\}$ vor mit

$$\mu = \frac{1}{6} \sum_{k=1}^6 k = \frac{21}{6} = 3.5$$

$$\sigma^2 = \frac{1}{6} \sum_{k=1}^6 k^2 - \mu^2 = \frac{1}{6} (1^2 + 2^2 + \dots + 6^2) - 3.5^2 \approx 2.917$$

$$\Rightarrow \sigma \approx 1.708$$

67.4 Die Binomialverteilung

Wir betrachten ein Experiment, das aus einer n -fachen Wiederholung eines Einzel-experiments

besteht, bei dem ein Ereignis A mit Wahrscheinlichkeit p eintritt. Ein solches Experiment heißt Bernoulli-Experiment. Wir setzen

$$X_i(\omega) = \begin{cases} 1 & \text{falls } A \text{ beim } i\text{-ten Versuch eintritt} \\ 0 & \text{sonst} \end{cases}$$

A_k bezeichnet das Ereignis, das im Gesamtexperiment gilt:

$$X(\omega) := \sum_{i=1}^n X_i(\omega) = k$$

Nach 62.4(c) (ungeordnete Stichprobe ohne Wiederholung) hat A_k insgesamt $\binom{n}{k}$ Realisierungen. Jede solche Realisierung tritt mit Wahrscheinlichkeit $p^k(1-p)^{n-k}$ auf. Die entsprechende Verteilung

$$b_{n,p}(k) := \binom{n}{k} p^k (1-p)^{n-k}$$

heißt Binomialverteilung mit Parametern n und p . Es gilt:

$$\begin{aligned} E(X_i) &= 1 \cdot p + 0 \cdot (1-p) \\ \Rightarrow E(X) &= \sum_{i=1}^n E(X_i) = np \\ V(X_i) &= E(X_i^2) - (E(X_i))^2 = 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 = p(1-p) \end{aligned}$$

Da die Einzelexperimente unabhängig sind, gilt:

$$V(X) = \sum_{i=1}^n V(X_i) = np(1-p)$$

67.5 Beispiel: Zufallsabhängigkeit sportlicher Resultate

Andreas und Bernd tragen ein Tischtennisturnier mit $n = 1, 3, 5, \dots, 2n + 1$ Spielen aus. Sieger ist, wer die meisten Einzelspiele gewinnt. Andreas gewinnt ein Einzelspiel mit Wahrscheinlichkeit $p = 0.6$. Wie groß sind die Siegeschancen für den schlechteren Spieler Bernd. Bernd siegt, wenn Andreas $S_n \leq m$ Erfolge erzielt. Die Wahrscheinlichkeit hierfür beträgt:

$$P(S_n \leq m) = b_{n,p}(0) + b_{n,p}(1) + \dots + b_{n,p}(m) = \sum_{k=0}^m \binom{n}{k} p^k (1-p)^{n-k}$$

- $n = 1$:

$$P(S_1 \leq 0) = \binom{1}{0} 0.6^0 \cdot 0.4^1 = 0.4$$

- $n = 3$:

$$P(S_3 \leq 1) = \binom{3}{0} 0.6^0 \cdot 0.4^3 + \binom{3}{1} 0.6^1 \cdot 0.4^2 \approx 0.352$$

- $n = 5$:

$$P(S_5 \leq 2) = \binom{5}{0} 0.6^0 \cdot 0.4^5 + \binom{5}{1} 0.6^1 \cdot 0.4^4 + \binom{5}{2} 0.6^2 \cdot 0.4^3 \approx 0.317$$

- $n = 7$:

$$P(S_7 \leq 3) \approx 0.290$$

- $n = 9$:

$$P(S_9 \leq 4) \approx 0.267$$

Es ist also gar nicht so unwahrscheinlich, dass der schlechtere Spieler das Turnier gewinnt.

67.6 Die Poisson-Verteilung

Für große n wird das Arbeiten mit der Binomialverteilung unhandlich. Ist p klein ($0 \leq p \leq 0.1$) gibt es eine gute Approximation: die Poisson-Verteilung zum Parameter λ :

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

Durch Umbenennen von „Erfolg“ und „Fehlschlag“ ist die Poisson-Verteilung auch für $0.9 \leq p \leq 1$ anwendbar.

Man kann zeigen, dass $p(k)$ für $\lambda = np$ die Binomialverteilung $b_{n,p}(k)$ approximiert. Ferner hat eine Poisson-verteilte Zufallsvariable den Erwartungswert λ und die Varianz λ . Generell beschreibt die Poisson-Verteilung Ereignisse, die im zeitlichen Verlauf zufällig und unabhängig voneinander auftreten, z.B.:

- atomarer Zerfall
- das Eintreffen von Bedienwünschen an einem Server
- Anrufe in einem Call-Center
- das Auftreten von Softwarefehlern in einem Programmsystem

67.7 Reales Beispiel: Der große Jubiläumstag

Genau in einem Jahr feiert ein großer Betrieb seinen 100. Geburtstag. Die Direktion beschließt, allen Kindern von Betriebsangehörigen, die an diesem Tag geboren werden, ein Sparkonton von 3000 Euro anzulegen. Da rund 730 Kinder pro Jahr geboren werden, erwartet man Auslagen von 6000 Euro. Um Zufallsschwankungen vorzubeugen, plant man 15000 Euro ein. Wie groß ist die Wahrscheinlichkeit, dass das Geld nicht reicht?

$n = 730$ Kinder / Jahr

$p = \frac{1}{365}$ Wahrscheinlichkeit, dass Geburtstag auf Jubiläumstag fällt.

$\Rightarrow \lambda = pn = 2.$

Das Geld reicht nicht, wenn $k \geq \frac{15000}{3000} = 5$ Kinder geboren werden.

$$\begin{aligned} p(k \geq 6) &= 1 - p(k \leq 5) = 1 - p(0) - p(1) - \dots - p(5) \\ &= 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} - \dots - \frac{2^5}{5!}e^{-2} \approx 0.0168 \end{aligned}$$

Die Wahrscheinlichkeit einer unangenehmen Zufallsüberraschung ist also gering.
Man rechnet nicht damit.

Anmerkung: Am Jubiläumstag wurden 36 Kinder geboren!

Die Direktion hat es also verstanden, ihre Betriebsangehörigen auch für außerbetriebliche Aktivitäten zu begeistern.

67.8 Die geometrische Verteilung

Eine diskrete Zufallsvariable X , die in einem Bernoulli-Experiment angeibt, wann ein Ereignis A mit Wahrscheinlichkeit $p = p(A)$ zum ersten Mal eintritt, heißt geometrisch verteilt mit Parameter p . Die entsprechende Verteilung lautet:

$$P_X(k) = p(1 - p)^{k-1}$$

Mann kann zeigen:

$$E(X) = \frac{1}{p}; \quad V(X) = \frac{1-p}{p^2}$$

Die geometrische Verteilung ist wichtig in der Informatik bei der Modellierung von Wartezeiten:

67.9 Beispiel:

Wie lange muss man beim Würfeln warten, bis die Augenzahl 4 erstmalig auftritt?

Mit $p = \frac{1}{6}$ beträgt der Erwartungswert

$$\frac{1}{p} = 6$$

und die Varianz

$$\frac{1-p}{p^2} = \frac{\frac{5}{6}}{\frac{1}{6^2}} = 30$$