Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine-Universität Düsseldorf

11. November 2010

1. Erwartungswert und Varianz
   - Erwartungswert
   - Varianz und Streuung
   - Rechenregeln
   - Binomialverteilung
   - Laplace-Verteilung
   - Geometrische Verteilung
   - Poissonverteilung

Erwartungswert: einleitendes Beispiel

- 100 unabhängige ja/nein Experimente
- Erfolgswahrscheinlichkeit im Einzelfall $p = 0.25$
- Wie viele Erfolge erwarten wir im Mittel?
- Antwort: 25 Erfolge im Mittel

Erwartungswert

- X eine diskrete Zufallsvariable
- Der *Erwartungswert* von X ist

$$E(X) = \sum_{k} P(X = k) \cdot k$$

- Die Summe läuft über alle Werte k von X

zurück zum Beispiel

- Im Beispiel $P(X = k) = B_{100,\,0.25}(k)$
- Laut Definition

$$E(X) = \sum_k k \cdot P(X = k) = \sum_{k=0}^{100} k \cdot B_{100,\,0.25}(k)$$

- Diese Zahl ist tatsächlich gleich 25
- Bei der ad-hoc Bestimmung vorhin wurden Rechenregeln benutzt

Beispiel: Würfel

Sei X die Augenzahl eines fairen Würfels.

$$\begin{aligned} E(X) &= \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 \\ &= \frac{21}{6} = \frac{7}{2} = 3.5 \end{aligned}$$

Im Mittel zeigt ein fairer Würfel 3.5 Augen

Beispiel: Glücksspiel

- X sei der Gewinn bei einem Glücksspiel
- $P(X = k)$ sei die Wahrscheinlichkeit, den Gewinn k zu erzielen
- Dann ist $E(X)$ der faire Einsatz

Spiel 77

Klasse	Ziffern	Gewinn	$P(X = k)$	$k \cdot P(X = k)$
I	7	177 777.00€	0.000 000 1	0.018€
II	6	77 777.00€	0.000 001 0	0.078€
III	5	7 777.00€	0.000 010 0	0.078€
IV	4	777.00€	0.000 100 0	0.078€
V	3	77.00€	0.001 000 0	0.077€
VI	2	17.00€	0.010 000 0	0.170€
VII	1	5.00€	0.100 000 0	0.500€

$$E(X) = 0.998€$$

Variante des Spiels 77

- Der Einsatz beträgt 2.50€. Bei welchen Hauptgewinn wäre das Spiel fair?
- Hauptgewinn sei J. Dann

$$\begin{aligned} E(X) &= 0.000\,000\,1 \cdot J + 0.000\,001 \cdot 77\,7777 \\ &+ 0.000\,01 \cdot 7\,777 + 0.000\,1 \cdot 777 \\ &+ 0.001 \cdot 77 + 0.01 \cdot 17 + 0.1 \cdot 5 \\ &= 0.000\,000\,1 \cdot J + 0.980 \end{aligned}$$

- Das soll gleich 2.50 sein. Also $0.000\,000\,1 \cdot J = 1.52$

$$J = \frac{1.52}{0.000\,000\,1} = 15\,200\,000$$

- Spiel 77 ist fair bei einem Hauptgewinn von 15.2 Millionen €

Alternative Berechnung des Erwartungswerts

Wenn $\omega_1, \omega_2, \ldots$ die Elementarereignisse sind, dann

$$E(X) = \sum_j X(\omega_j) \cdot P(\omega_j)$$

Die Summe läuft über alle möglichen Elementarereignisse.

In dieser Darstellung sieht man die Verwandschaft zum arithmetischen Mittel.

Varianz und Streuung

Die *Varianz* einer Zufallsvariablen X ist definiert als

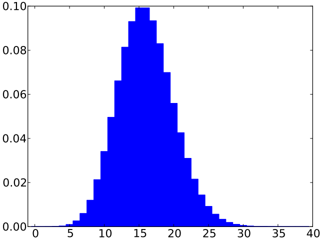
$$Var(X) = \sum_{k=0}^{\infty} (k - \mu)^2 \cdot P(X = k)$$

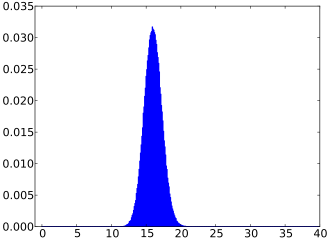
wobei $\mu = E(X)$.

Die *Standardabweichung* oder *Streuung* von X ist definiert als die Wurzel aus der Varianz

$$\sigma = \sqrt{Var(X)}$$

Gleicher Erwartungswert, unterschiedliche Streuung





Links: Verteilung X_1 mit $E(X_1) = 16$ und $\sigma = 4$

Rechts: Verteilung X_2 mit $E(X_2) = 16$ und $\sigma = 1.26$

Alternative Berechnung der Varianz

Wenn $\omega_1, \omega_2, \ldots$ die Elementarereignisse sind, dann

$$Var(X) = \sum_j (X(\omega_j) - E(X))^2 \cdot P(\omega_j)$$

Die Summe läuft über alle möglichen Elementarereignisse.

Modell vs. Datensatz

Datensatz	Modell
arithmetisches Mittel	Erwartungswert
empirische Varianz	Varianz
Stichprobenstreuung	Streuung

Rechenregeln

Rechenregeln für den Erwartungswert

- Für jede Zahl c und jede Zufallsvariable X ist
  $E(c \cdot X) = c \cdot E(X)$
- Für Zufallsvariablen $X_1, \ldots, X_n$ ist
  $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$

Rechenregeln für die Varianz

- Für jede Zahl a und jede Zufallsvariable X gilt
  $Var(a + X) = Var(X)$
- Für Zahl c und jede Zufallsvariable X gilt
  $Var(c \cdot X) = c^2 \cdot Var(X)$
- Für jede Zufallsvariable X gilt $Var(X) = E(X^2) - E(X)^2$

Unabhängigkeit von Zufallsvariablen

Zwei diskrete Zufallsvariable X und Y sind *stochastisch unabhängig*, wenn für alle möglichen Werte k und m

$$P(X = k, Y = m) = P(X = k) \cdot P(Y = m)$$

Zusätzliche Rechenregeln für unabhängige Zufallsvariable

Produktformel für den Erwartungswert: X und Y seien **unabhängige** Zufallsvariable. Dann

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

Summenformel für die Varianz: X und Y seien **unabhängige** Zufallsvariable. Dann

$$Var(X + Y) = Var(X) + Var(Y)$$

Einzelnes ja/nein-Experiment

Zufallsvariable X nimmt mit Wahrscheinlichkeit p den Wert 1 und mit Wahrscheinlichkeit $1-p$ den Wert 0 an

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{1} k \cdot P(X=k) \\
&= 0 \cdot (1-p) + 1 \cdot p \\
&= p \\
Var(X) &= \sum_{k=0}^{1} (k - E(X))^2 \cdot P(X=k) \\
&= p^2 \cdot (1-p) + (1-p)^2 \cdot p \\
&= p \cdot (1-p) \cdot (p + (1-p)) \\
&= p \cdot (1-p)
\end{aligned}
$$

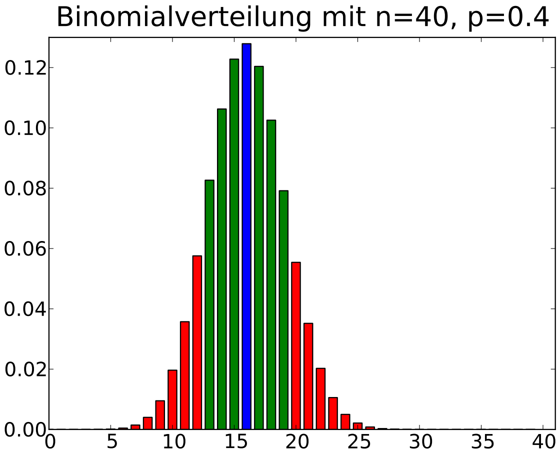
Erwartungswert der Binomialverteilung

- $X_1, X_2, \ldots, X_n$ seien Zufallsvariable
- jedes X_j nimmt mit Wahrscheinlichkeit p den Wert 1 und mit Wahrscheinlichkeit $1 - p$ den Wert 0 an
- Interpretation $X_j = 1$ ist "Erfolg"
- $Y = X_1 + X_2 + \cdots + X_n$ ist die Anzahl der Erfolge
- Y ist verteilt gemäß $B_{n,p}$
- $E(X_1) = E(X_2) = \cdots = E(X_n) = p$
- also $E(Y) = n \cdot p$
- Der Erwartungswert einer $B_{n,p}$-verteilten Zufallsvariable beträgt $n \cdot p$
- Dieses Argument steht hinter der früheren Überlegung

Varianz der Binomialverteilung

- $X_1, X_2, \ldots, X_n$ seien **unabhängige** Zufallsvariable
- jedes X_j nimmt mit Wahrscheinlichkeit p den Wert 1 und mit Wahrscheinlichkeit $1 - p$ den Wert 0 an
- $Y = X_1 + X_2 + \cdots + X_n$ ist verteilt gemäß $B_{n,p}$
- $Var(X_1) = Var(X_2) = \cdots = Var(X_n) = p \cdot (1 - p)$
- also $Var(Y) = n \cdot p \cdot (1 - p)$
- Die Varianz einer $B_{n,p}$-verteilten Zufallsvariable beträgt $n \cdot p \cdot (1 - p)$

Stabdiagramm der Binomialverteilung



Der Erwartungswert ist blau, alles innerhalb der Standardabweichung ist grün

Erwartungswert und Varianz: Beispiel Laplace-Verteilung

Die Zufallsvariable X sei Laplace-verteilt auf $\{1, 2, \ldots, N\}$. Dann

$$E(X) = \frac{N+1}{2}$$

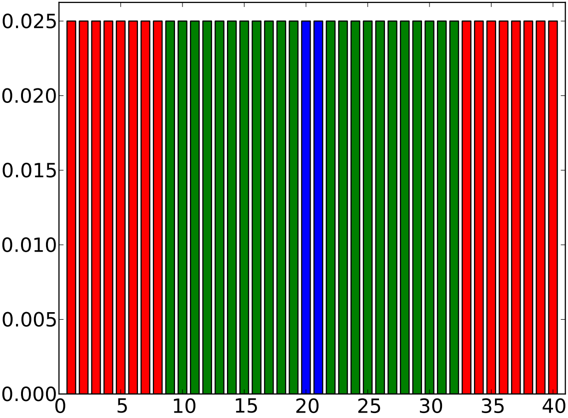
und

$$Var(X) = \frac{N^2-1}{12}$$

Für den fairen Würfel gelten also

$$E(X) = \frac{7}{2} = 3.5$$
$$Var(X) = \frac{35}{12} = 2.917$$
$$\sigma = \sqrt{\frac{35}{12}} = 1.708$$

Stabdiagramm Laplace-Verteilung



Laplace-Verteilung auf $\{1, 2, \ldots, 40\}$. Der Erwartungswert ist blau, alles innerhalb der Standardabweichung ist grün.

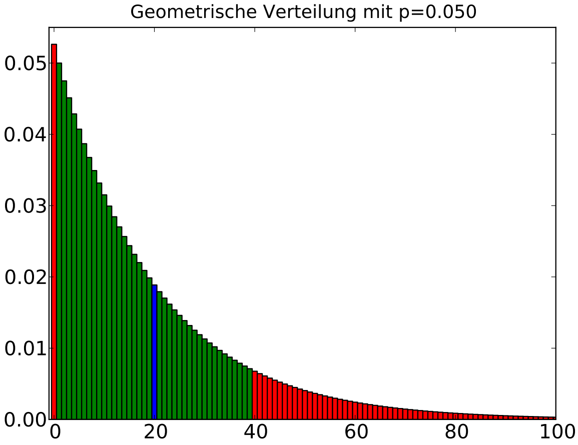
Erwartungswert und Varianz für die geometrische Verteilung

Die Zufallsvariable X sei G_p-verteilt. Dann

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

Stabdiagramm geometrische Verteilung



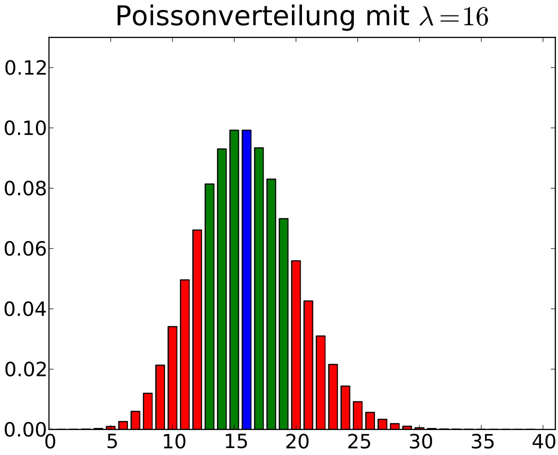
Geometrische Verteilung $G_{0.05}$. Der Erwartungswert ist blau, alles innerhalb der Standardabweichung ist grün.

Erwartungswert und Varianz für die Poissonverteilung

Die Zufallsvariable X sei P_λ-verteilt. Dann

$$E(X) = \lambda$$
$$Var(X) = \lambda$$

Stabdiagramm Poissonverteilung



Poissonverteilung P_{16}. Der Erwartungswert ist blau, alles innerhalb der Standardabweichung ist grün.