Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun
http://blog.ruediger-braun.net

Heinrich-Heine-Universität Düsseldorf

08. Januar 2014

Binomialverteilung

- n unabhängige Wiederholungen eines ja/nein-Experiments
- das Ergebnis "ja" bezeichnet man als "Erfolg"
- die Wahrscheinlichkeit von "ja" sei p, genannt "Erfolgswahrscheinlichkeit im Einzelfall"
- $B_{n,p}(k)$ ist die Wahrscheinlichkeit von genau k Erfolgen

$$B_{n,p}(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

- Beispiel:

$$B_{10,\,1/6}(2) = \binom{10}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^8 = 45 \cdot 0.02778 \cdot 0.2326 = 0.2907$$

- $B_{n,p}(n) = p^n$ und $B_{n,p}(0) = (1-p)^n$

Beispiel zur Binomialverteilung: L-Bakterien

- Chiralität: Manche Moleküle kommen bei gleicher chemischer Zusammensetzung in zwei verschiedenen geometrischen Formen vor. Man bezeichnet dann eine von beiden als "rechtsdrehend", die andere als "linksdrehend"
- Von einem bestimmten Bodenbakterium gibt es zwei Varianten, je nachdem ob ein bestimmtes Molekül in der linksdrehenden oder der rechtsdrehenden Version bevorzugt wird. Wir nennen die beiden Varianten L-Bakterium und R-Bakterium.
- In ungestörtem Boden befinden sich 75% L-Bakterien und 25% R-Bakterien.

Beispiel zur Binomialverteilung: L-Bakterien

- Die Bakterien werden einem Pestizid ausgesetzt. Nach einigen Generationen werden die Nachkommen getestet, ob sie rechts- oder linksdrehende Varianten aufnehme
- 13 der 27 Nachkommen sind R-Bakterien. Ist dieses Ergebnis so ungewöhnlich, dass man einen Einfluss des Pestizids vermuten muss?

L-Bakterien: Fortsetzung

- Binomialverteilung $B_{n,p}(k)$

$$\begin{array}{rcr} \text{Erfolg} & : & L\text{-Bakterium} \\ p & : & 0.75 \\ n & : & 27 \\ k & : & 14 \end{array}$$

- Mit Wahrscheinlichkeit

$$B_{27,\,0.75}(14) = \binom{27}{14} \cdot 0.75^{14} \cdot 0.25^{13} = 0.005326$$

sind 14 von 27 Bakterien L-Bakterien

- Das ist aber nicht die richtige Frage.

L-Bakterien: Fortsetzung

- Wenn weniger als 14 der Bakterien L-Bakterien gewesen wären, hätten wir das noch ungewöhnlicher gefunden
- Wir suchen also die Wahrscheinlichkeit, dass von 27 Bakterien höchstens 14 L-Bakterien sind,
- d. h. wir suchen
$$\sum_{k=0}^{14} B_{27,\,0.75}(k)$$
- Diese Zahlen findet man in Tabellen der kumulierten Binomialverteilung

Tabelle der Werte $\sum_{k=0}^{r} B_{n,p}(k)$ für $n = 27$

r	p	0.75	0.76	0.77	0.78	0.79
9	0.	00001				
10		00003	00002	00001	00001	
11		00016	00010	00006	00003	00002
12		00067	00042	00026	00015	00009
13		00245	00161	00103	00065	00039
14		00778	00538	00364	00240	00153
15		02162	01573	01119	00777	00526
16		05278	04031	03016	02208	01577
17		11325	09067	07126	05488	04136
18		21405	17927	14769	11951	09483
19		35729	31217	26889	22804	19012
20		52917	48050	43120	38195	33350
21		70105	65819	61232	56385	51330
22		84168	81165	77770	73973	69777
23		93340	91729	89806	87530	84864
24		97926	97305	96521	95540	94322
25		99577	99423	99219	98948	98592
26		99958	99939	99914	99878	99828

L-Bakterien: Fortsetzung

Mit Wahrscheinlichkeit 0.00778 sind von 27 Bakterien höchstens 14 L-Bakterien.

Das legt den Verdacht nahe, dass das Pestizid hauptsächlich L-Bakterien angreift.

Geometrische Verteilung

Die geometrische Verteilung modelliert Wartezeiten in getakteten Modellen

Geometrische Verteilung, Beispiel

Wurf eines fairen Würfels:

- Erfolgswahrscheinlichkeit im Einzelfall $p = \frac{1}{6}$
- Die Wahrscheinlichkeit, dass im ersten Wurf eine Sechs fällt, ist
$$\frac{1}{6} = 0.1667$$
- Die Wahrscheinlichkeit, dass die erste Sechs im zweiten Wurf fällt, ist
$$\frac{5}{6} \cdot \frac{1}{6} = 0.1389$$

Beispiel, Fortsetzung

Wurf eines fairen Würfels:

- Die Wahrscheinlichkeit, dass die erste Sechs im dritten Wurf fällt, ist
$$\left(\frac{5}{6}\right)^2 \cdot \frac{1}{6} = 0.1157$$
- Die Wahrscheinlichkeit, dass die erste Sechs im k-ten Wurf fällt, ist
$$\left(\frac{5}{6}\right)^{k-1} \cdot \frac{1}{6}$$

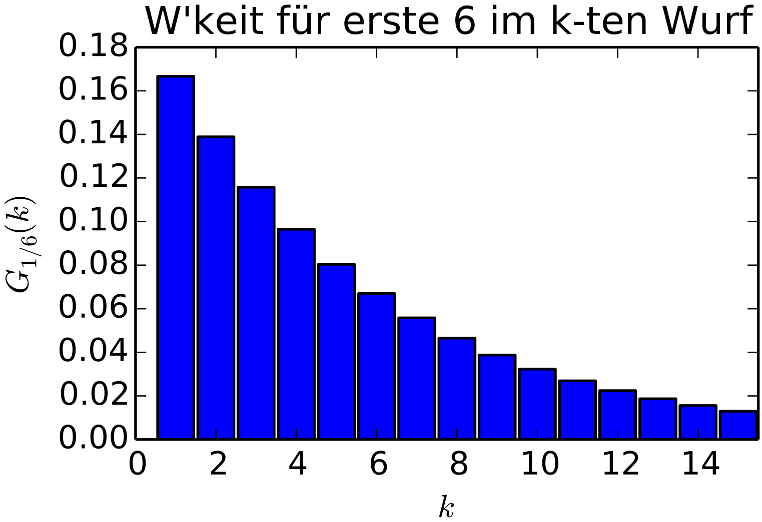
Geometrische Verteilung: Definition

Die *geometrische Verteilung* zum Parameter p wird gegeben durch

$$G_p(k) = (1-p)^{k-1} \cdot p$$

In einem ja/nein-Experiments, bei dem die Erfolgswahrscheinlichkeit im Einzelfall gleich p ist, gebe die Zufallsvariable X die Nummer des Zugs mit dem ersten Erfolg an. Dann ist X verteilt gemäß G_p.

Stabdiagramm von $G_{1/6}$



Geometrische Verteilung: Beispiel

In einer bestimmten Bucht beträgt im Juli die Wahrscheinlichkeit einer Walbeobachtung an einem zufällig ausgewählten Tag 15%

- Mit welcher Wahrscheinlichkeit sieht man 7 Tage lang keinen einzigen Wal?
- Wie groß ist die Wahrscheinlichkeit, dass man seinen ersten Wal am siebten Tag sieht?

Walbeobachtung, Fortsetzung

Erste Frage:

Mit welcher Wahrscheinlichkeit sieht man 7 Tage lang keinen einzigen Wal?

- 7 unabhängige ja/nein-Experimente: Binomialverteilung
- Erfolgswahrscheinlichkeit im Einzelfall $p = 0.15$
- gesucht: Wahrscheinlichkeit von 0 Erfolgen

$$B_{7\,0.15}(0) = \binom{7}{0} p^0 (1-p)^7 = 0.85^7 = 0.3206$$

- Die Wahrscheinlichkeit, dass 7 Tage lang kein einziger Wal beobachtet wird, beträgt 32%

Walbeobachtung, Fortsetzung

Zweite Frage

Wie groß ist die Wahrscheinlichkeit, dass man seinen ersten Wal am siebten Tag sieht?

- Wartezeit in getaktetem Modell: Geometrische Verteilung
- gesucht: Wahscheinlichkeit, dass erster Erfolgt im siebten Versuch eintritt, wenn Erfolgswahrscheinlichkeit im Einzelfall $p = 0.15$

$$G_{0.15}(7) = 0.85^6 \cdot 0.15 = 0.0566$$

- Die Wahrscheinlichkeit, dass die erste Walbeobachtung am siebten Tag statt findet, beträgt 5.7%

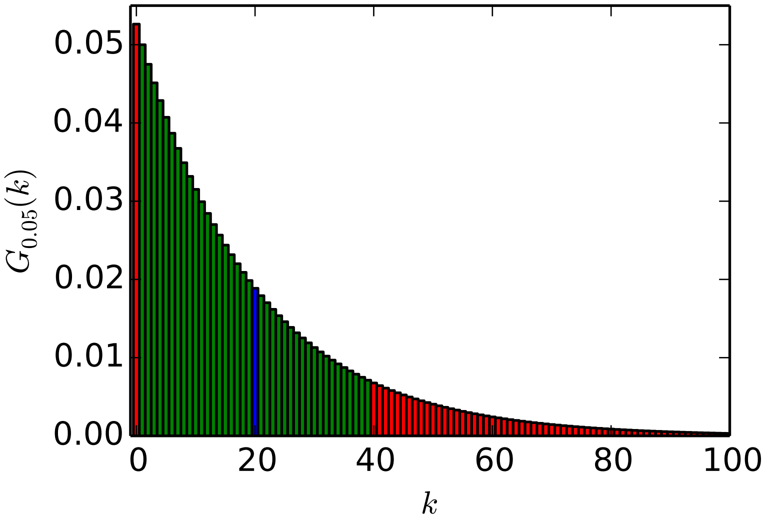
Erwartungswert und Varianz für die geometrische Verteilung

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.

Beispiel: Walbeobachtung

In einer bestimmten Bucht beträgt im Juli die Wahrscheinlichkeit einer Walbeobachtung an einem zufällig ausgewählte Tag 15%

- Wie viele Tage mit Walbeobachtung gibt es im Mittel in einer Woche?
- An welchem Tag sieht man im Mittel den ersten Wal?

Walbeobachtung: Erste Frage

Wie viele Tage mit Walbeobachtung gibt es im Mittel in einer Woche?

- Die Zufallsvariable X zählt die Tage mit Walbeobachtung in einer Woche
- X ist $B_{n,\,p}$-verteilt für $n = 7$ und $p = 0.15$
- Also $E(X) = n \cdot p = 7 \cdot 0.15 = 1.05$
- Im Mittel gibt es pro Woche 1.05 Tage mit Walbeobachtung

Walbeobachtung: Zweite Frage

An welchem Tag sieht man im Mittel den ersten Wal?

- Die Zufallsvariable Y gibt den ersten Tag mit Walbeobachtung an.
- Y ist G_p-verteilt für $p = 0.15$
- Also $E(Y) = \dfrac{1}{p} = \dfrac{1}{0.15} = 6.667$
- Im Mittel ist der siebte Tag der erste mit Walbeobachtung

Poissonverteilung

Die Poissonverteilung modelliert seltene Ereignisse

Selten bedeutet, dass sich die Einzelereignisse nicht beeinflussen

Es sei $\lambda > 0$. Die Poissonverteilung zum Parameter λ ist definiert durch

$$P_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

Poissonverteilung, Bedeutung

Unter den folgenden Voraussetzungen ist eine Zufallsvariable X poissonverteilt zum Parameter λ:

- X zählt das Auftreten eines Ereignisses pro Zähleinheit
- Im Mittel treten λ Ereignisse pro Zähleinheit auf
- Die Ereignisse beeinflussen sich nicht gegenseitig

Beispiel: α-Strahlung

Versuchsaufbau:

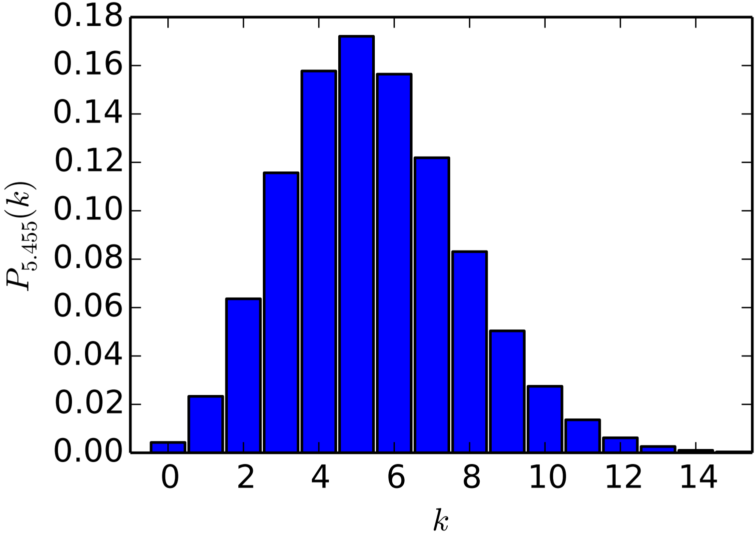
- Wir beobachten die Einschläge von α-Teilchen in einer Zelle
- Im Mittel gebe es einen Einschlag alle 11 Minuten, also 5.455 Einschläge pro Stunde.
- Welcher Prozentsatz der Zellen übersteht eine Stunde ohne Einschlag?

Es ist $\lambda = 5.455$, und wir suchen den Wert von P_λ für $k = 0$

$$P_\lambda(0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-5.455} = 0.004277$$

Nur ungefähr jede 230-te Zelle bleibt unbeschädigt

Stabdiagramm von $P_{5.455}$

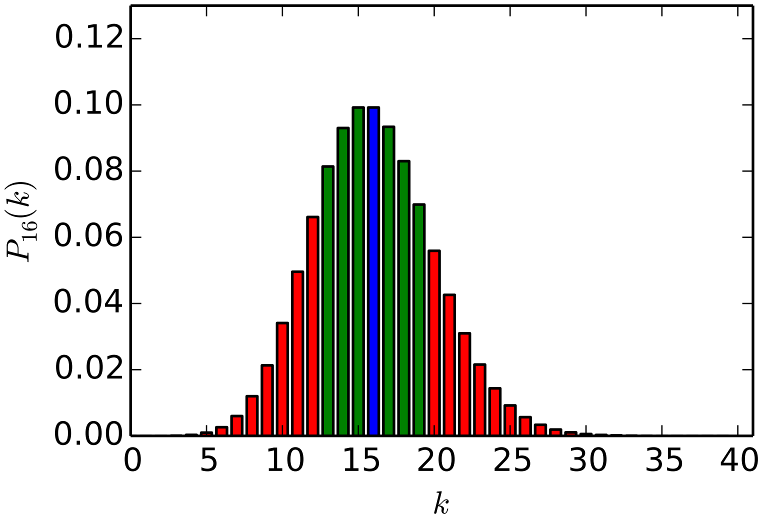


Erwartungswert und Varianz für die Poissonverteilung

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.