Mathematik II für Biologen

Binomialverteilung & Binomialtest

Stefan Keppeler

28. Mai 2008

Kombinatorik
- Permutationen
- Urnenmodelle
- Binomialkoeffizient

Binomialverteilung
- Motivation
- $\mathrm{Bin}(n, p)$
- Histogramme

Binomialtest
- Beispiel
- Faustregeln
- Vorzeichentest

- Anzahl möglicher Anordungen von n (unterschiedlichen) Elementen:

$$n(n-1)\cdots 1 = n!$$

Beispiel: $n = 3$ Kugeln, ●●●, lassen sich auf $3! = 6$ verschiedene Möglichkeiten anordnen:
(●●●) (●●●) (●●●) (●●●) (●●●) (●●●)

- Anzahl möglicher Anordungen von n (unterschiedlichen) Elementen:

$$n(n-1)\cdots 1 = n!$$

Beispiel: $n = 3$ Kugeln, ●●●, lassen sich auf $3! = 6$ verschiedene Möglichkeiten anordnen:
(●●●) (●●●) (●●●) (●●●) (●●●) (●●●)

- Anzahl möglicher Anordungen von n Elementen, die in j Gruppen mit jeweils k_1, $k_2, \ldots k_j$ gleichen Elementen vorkommen ($k_1 + k_2 + \ldots + k_j = n$):

$$\frac{n!}{k_1!\,k_2!\cdots k_j!}$$

Beispiel: $k_1 = 2$ rote Kugeln und $k_2 = 1$ blaue Kugel, ●●●, lassen sich auf $\frac{3!}{2!\,1!} = 3$ verschiedene Möglichkeiten anordnen:
(●●●) (●●●) (●●●)

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge		
ohne Beachtung der Reihenfolge		

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k	
ohne Beachtung der Reihenfolge		

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
ohne Beachtung der Reihenfolge		

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
ohne Beachtung der Reihenfolge		$\frac{n!}{(n-k)!\,k!} =: \binom{n}{k}$

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
ohne Beachtung der Reihenfolge	$\binom{n+k-1}{k}$	$\frac{n!}{(n-k)!\,k!} =: \binom{n}{k}$

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse:

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k Würfelergebnisse: $6^3 = 216$	$\dfrac{n!}{(n-k)!}$ zweifarbige Zwerge: $6 \cdot 5 = 30$
ohne Beachtung der Reihenfolge	$\dbinom{n+k-1}{k}$ Kartenzahl: $\binom{6+3-1}{3} = 56$	$\dfrac{n!}{(n-k)!\,k!} =: \dbinom{n}{k}$ dreifarbige Zwerge: $\binom{6}{3} = 20$

Beispiel: Würfelzwerge (ÜA 19)

Dabei heißt

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

Binomialkoeffizient, lies "n *über* k" oder "k *aus* n".

Beispiele:

- Betrachte ein Experiment mit zwei möglichen Ergebnissen.
- Nenne den einen Ausgang *"Erfolg"*, den anderen *"Misserfolg"*.
- Die Erfolgswahrscheinlichkeit sei $p \in [0, 1]$.
- Führe das Experiment n mal unabhängig durch (keine gegenseitige Beeinflussung der Einzelergebnisse).

Wie groß ist die Wahrscheinlichkeit für k Erfolge? ($0 \leq k \leq n$)

- Betrachte ein Experiment mit zwei möglichen Ergebnissen.
- Nenne den einen Ausgang *"Erfolg"*, den anderen *"Misserfolg"*.
- Die Erfolgswahrscheinlichkeit sei $p \in [0, 1]$.
- Führe das Experiment n mal unabhängig durch (keine gegenseitige Beeinflussung der Einzelergebnisse).

Wie groß ist die Wahrscheinlichkeit für k Erfolge? ($0 \leq k \leq n$)

Beispiele: (jeweils n Würfe)

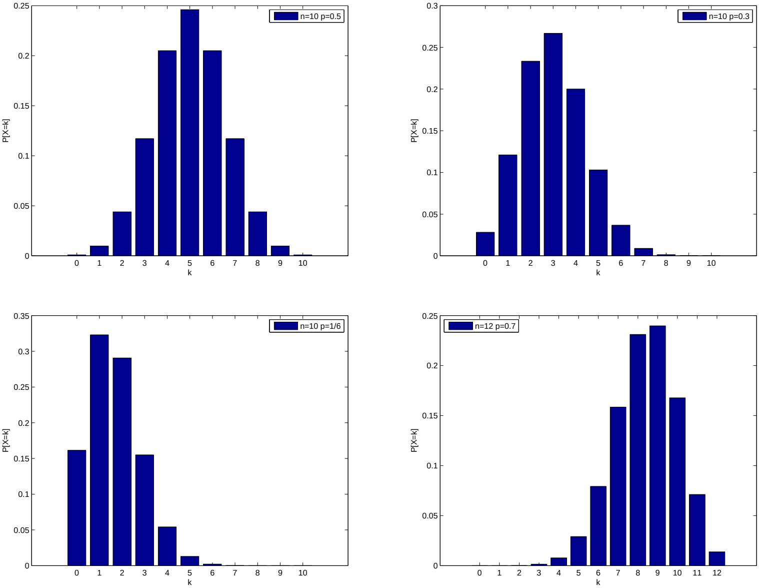
- faire Münze, $X = \#$ Zahl, $p = \frac{1}{2}$
- unfaire Münze, $X = \#$ Zahl, z.B. $p = 0{,}3$
- fairer Würfel, $X = \#$ Sechsen, $p = \frac{1}{6}$

Satz: Bei einer Serie von n unabhängigen Experimenten mit Erfolgswahrscheinlichkeit p, $0 \leq p \leq 1$, ist die Verteilung der Anzahl X der Erfolge gegeben durch

$$P[X{=}k] = \binom{n}{k} p^k \, (1-p)^{n-k} \, .$$

Man sagt X ist binomialverteilt mit Parametern n und p, Schreibweise $X \sim \mathrm{Bin}(n,p)$.

Beweis:


n=10 p=0.5
n=10 p=0.3
n=10 p=1/6
n=12 p=0.7
P[X=k]
k

Beispiel: Spermasexing

- Rindersperma wird optisch nach den y-Chromosomen sortiert.
- $n = 12$ Kühe werden damit besamt.
- *"Erfolg"*: weibliches Kalb – $♀$
  "Misserfolg": männliches (oder gar kein) Kalb – $♂$
- p: Wahrscheinlichkeit, dass (einmalige) Besamung zu $♀$ führt (Ziel: $p \approx 1$, denn $♀$ gibt Milch, $♂$ nicht)
- $X = \#$ gelungene Experimente $= \# ♀$

Beispiel: Spermasexing

- Rindersperma wird optisch nach den y-Chromosomen sortiert.
- $n = 12$ Kühe werden damit besamt.
- *"Erfolg"*: weibliches Kalb – ♀
  "Misserfolg": männliches (oder gar kein) Kalb – ♂
- p: Wahrscheinlichkeit, dass (einmalige) Besamung zu ♀ führt (Ziel: $p \approx 1$, denn ♀ gibt Milch, ♂ nicht)
- $X = \#$ gelungene Experimente $= \#$ ♀

Anbieter der Methode behauptet $p > 0{,}7$ und will dies durch einen statistischen Test beweisen.

- Wir erhalten $X = 11$ (d.h. $11/12 \approx 92\%$ Erfolsquote)

Beispiel: Spermasexing

- Rindersperma wird optisch nach den y-Chromosomen sortiert.
- $n = 12$ Kühe werden damit besamt.
- *"Erfolg"*: weibliches Kalb – ♀
  "Misserfolg": männliches (oder gar kein) Kalb – ♂
- p: Wahrscheinlichkeit, dass (einmalige) Besamung zu ♀ führt (Ziel: $p \approx 1$, denn ♀ gibt Milch, ♂ nicht)
- $X = \#$ gelungene Experimente $= \#$ ♀

Anbieter der Methode behauptet $p > 0{,}7$ und will dies durch einen statistischen Test beweisen.

- Wir erhalten $X = 11$ (d.h. $11/12 \approx 92\%$ Erfolsquote)

Test: (verwende in Testschritt 4 statt Simulation nun Wahrscheinlichkeitsrechnung)

Faustregeln für den Binomialtest

Falls $n\,p\,(1-p)$ groß ist (>9), so ist für $\alpha = 5\%$

- der Annahmebereich K^C zu
  $H_0 : p = p_0$ und $H_A : p \neq p_0$ etwa

$$K^C \approx np_0 \pm 2\sqrt{np_0(1-p_0)}$$

und

- das 95%-Vertrauensintervall für den wahren Wert von p

$$\approx \left[\frac{X}{n} - 2\sqrt{\frac{1}{n}\frac{X}{n}\left(1-\frac{X}{n}\right)}, \frac{X}{n} + 2\sqrt{\frac{1}{n}\frac{X}{n}\left(1-\frac{X}{n}\right)}\right].$$

Begründung (später): Wird näherungsweise Normalverteilung

Spezialfall des Binomialtests mit $p_0 = \frac{1}{2}$: Vorzeichentest

Beispiel:

- Eine Waage gelte als geeicht, falls sie
  - mit Wahrscheinlichkeit 50% einen zu großen und
  - mit Wahrscheinlichkeit 50% einen zu kleinen Wert anzeigt.
- Sollwert $20\,\mathrm{kg}$, $n = 10$ Messungen ergeben (in kg):
  $20{,}1 \quad 20{,}3 \quad 20{,}9 \quad 19{,}2 \quad 20{,}8 \quad 20{,}1 \quad 20{,}2 \quad 20{,}4 \quad 20{,}2 \quad 20{,}3$
- Test