Testen von Hypothesen

Erfurt 5.3.2015

Dr. Hubert Langlotz hlanglotz@gmx.de

Thillm

Thüringer Institut für Lehrerfortbildung, Lehrplanentwicklung und Medien

T³ DEUTSCHLAND

14. Thüringer Regionaltagung T³/MNU

Jena

26. September 2015

9:00 – 15:30 Uhr

Mathematisch-naturwissenschaftlicher Unterricht mit CAS an allen Thüringer Gymnasien, Gesamtschulen, Berufsbildenden Schulen mit gymnasialer Oberstufe und Kollegs

Was sagt der Lehrplan

Für alle:

Der Schüler kann

- exemplarisch statistische Erhebungen planen und beurteilen,
- zweiseitige Signifikanztests für binomialverteilte Zufallsgrößen durchführen und interpretieren,

Thema 2 (Wahlthema)

-Hypothesentests (auch einseitige Signifikanztests und Alternativtests) für binomial- und normalverteilte Zufallsgrößen durchführen und interpretieren,

-Unsicherheit der Ergebnisse von Hypothesentests begründen



Ist das Werfen einer Münze eigentlich ein faires Spiel?

Der entscheidende Schritt beim Signifikanztest liegt weniger in der rechnerischen Durchführung des Tests als in der *Wahl der Nullhypothese* und ihrer Gegenhypothese. Ein Unterricht, der sich auf den Formalismus eines Signifikanztests, seine rein mathematische Ausprägung beschränkt, wird dem stochastischen Aspekt nicht gerecht. Da man nicht anders zeigen kann, dass eine Hypothese A gegenüber einer Hypothese B den Vorzug verdient, als dass man B statistisch begründet zurückweist, wird die Nullhypothese aufgestellt mit der Absicht, sie nach Möglichkeit zu verwerfen. Dabei darf man die Möglichkeit einer Fehlentscheidung nicht außer Acht lassen, insbesondere wenn sie schwerwiegende Folgen haben könnte.

August Schmid „Verständnis lehren“ , S. 322, Klett

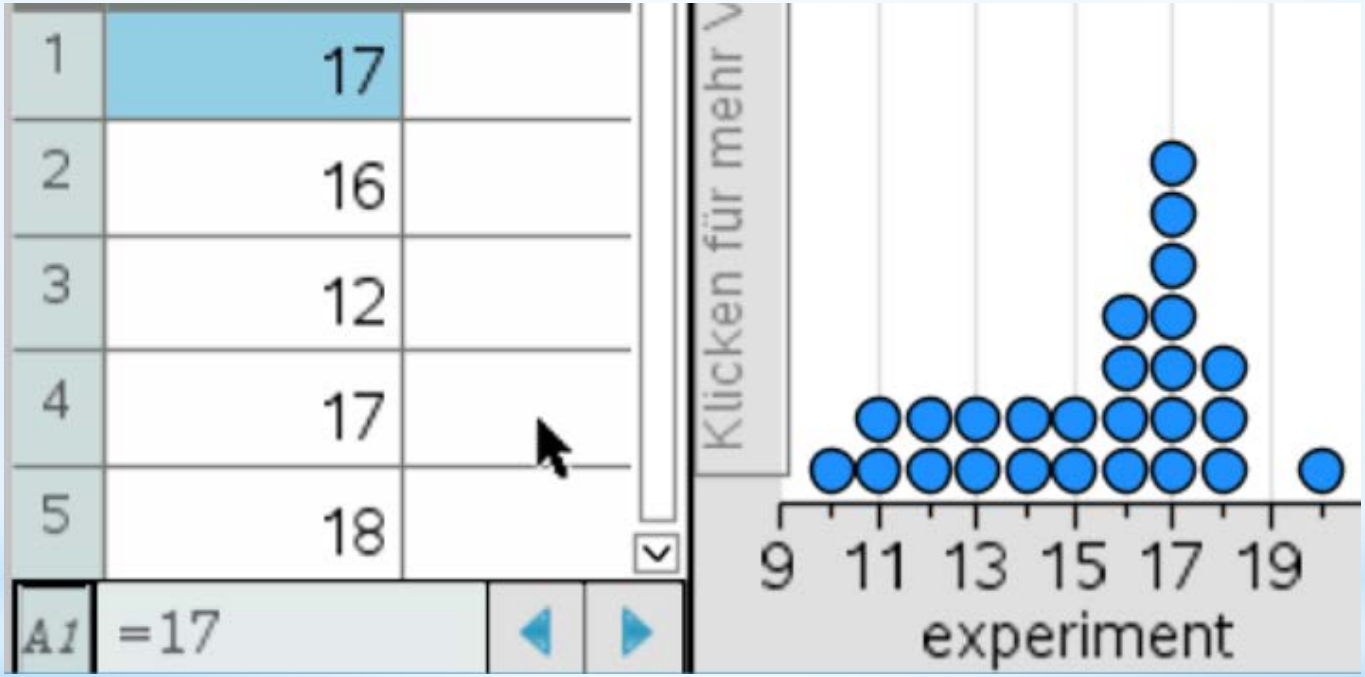
Experiment zum Münztest

Wann würden wir meinen, dass eine Münze nicht fair ist?

Jeder Schüler wirft die 1Cent-Münze 30mal.

Wir zählen, wie häufig Kopf erscheint.

Ergebnisse des Experiments



Alle Münzen waren neu!

Kernidee des Hypothesentestens

Weichen die tatsächlichen Daten von der bei der Gültigkeit der Nullhypothese zu erwartenden Situation ”überzufällig“ stark ab, d.h. so stark, dass man die Abweichung nicht mehr nur der Zufallsstreuung zuschreiben kann?

Nur in diesem Fall ist die Nullhypothese abzulehnen.

Idee:

-kleine Abweichung nur Zufallsstreuung

-große Abweichung Zufallsstreuung + inhaltlicher Unterschied
Nullhypothese ablehnen

Dr. Gero Walter LMU München
http://www.statistik.lmu.de/~walter/lehre/Stat2Soz_09/material/Stat2Soz_09_Skript-Teil3.pdf

Entscheidungsregel

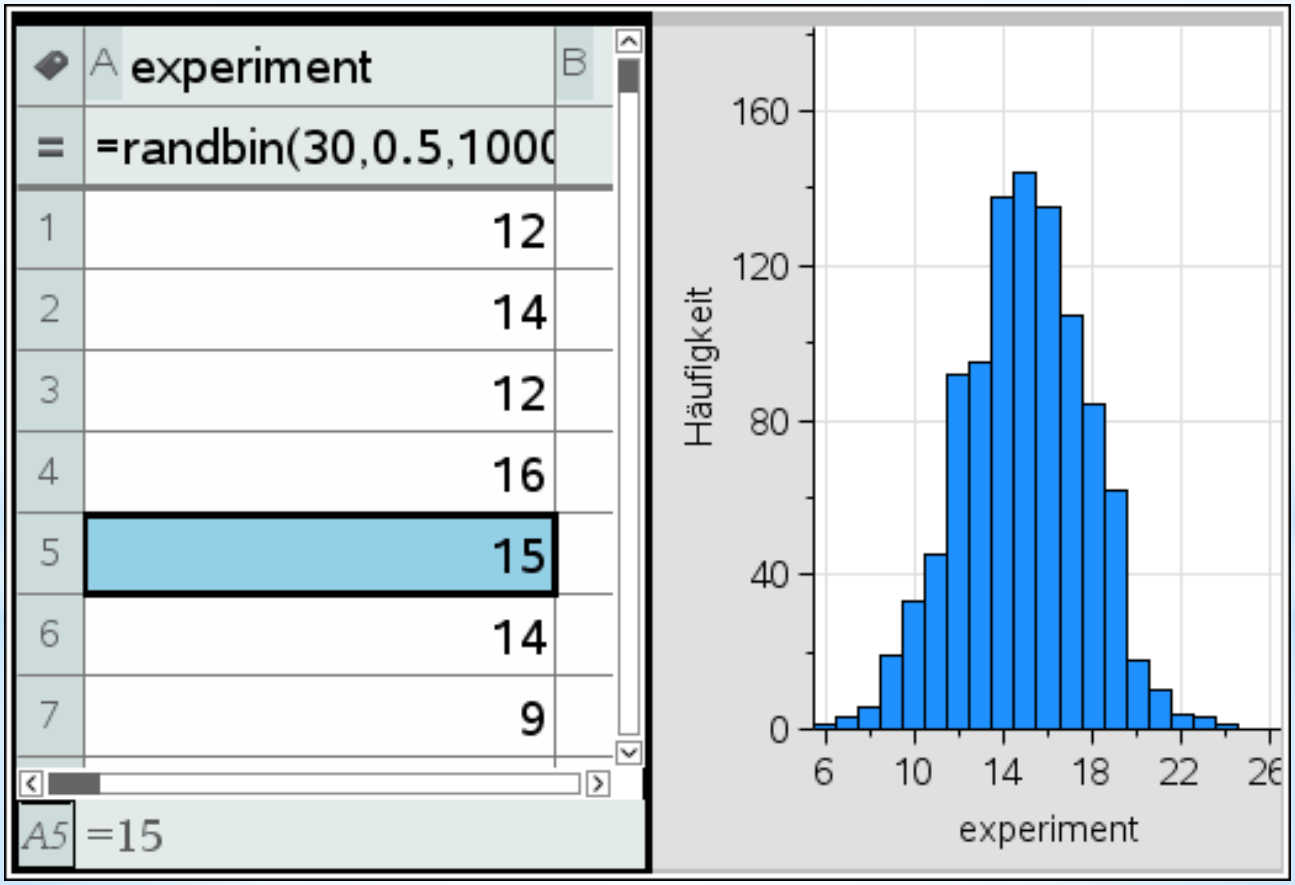
Wann sollten wir die Hypothese „Die Münze ist fair, ablehnen?

Bei 30 Würfen:

Wir würden die Münze noch als fair ansehen, wenn die Anzahl von Kopf zwischen … und …

liegt.

Simulation mit einem digitalen Werkzeug



Zweiseitiger Signifikanztest allgemeines Vorgehen

Variante 1

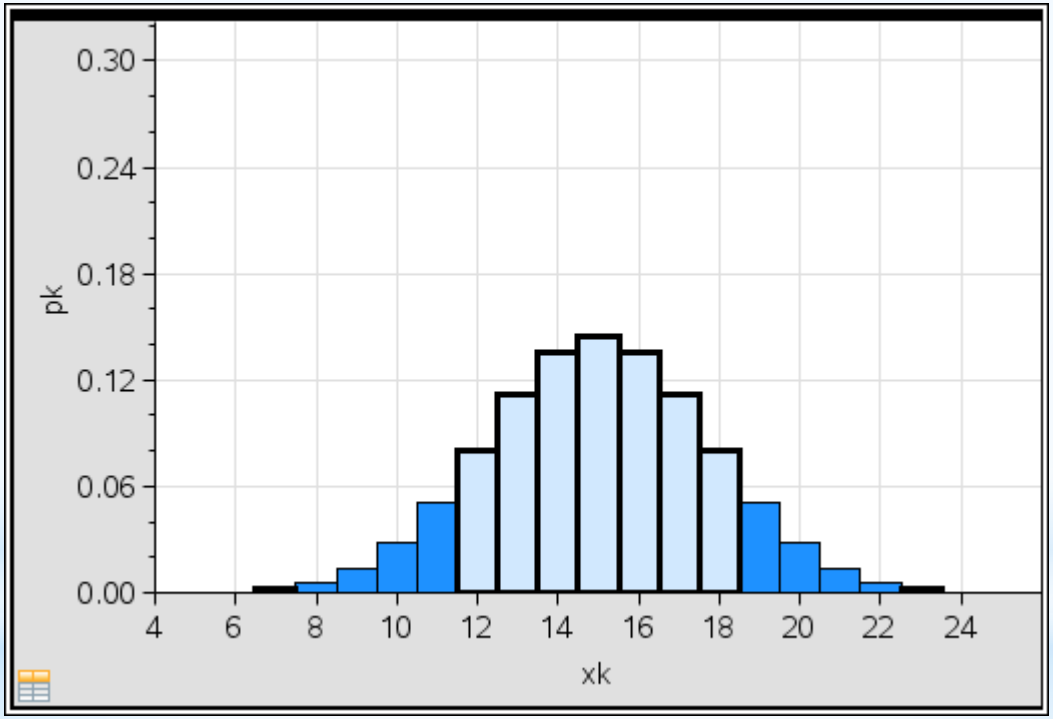
1. Festlegung von H0, H1

2. Nicht-Ablehnungs- und Ablehnungsbereich sind vorgegeben (werden gewählt)

3. Fehler 1. Art kann berechnet werden.

$$\text{binomCdf}(30,0.5,0,11)+\text{binomCdf}(30,0.5,19,30)$$

$$0.200488$$



Zweiseitiger Signifikanztest allgemeines Vorgehen

Variante 2

1. Festlegung von H_0, H_1

2. Signifikanzniveau vorgeben

eine bereits vor dem Test festgelegte obere Grenze α für die Wahrscheinlichkeiten des Fehlers 1. Art

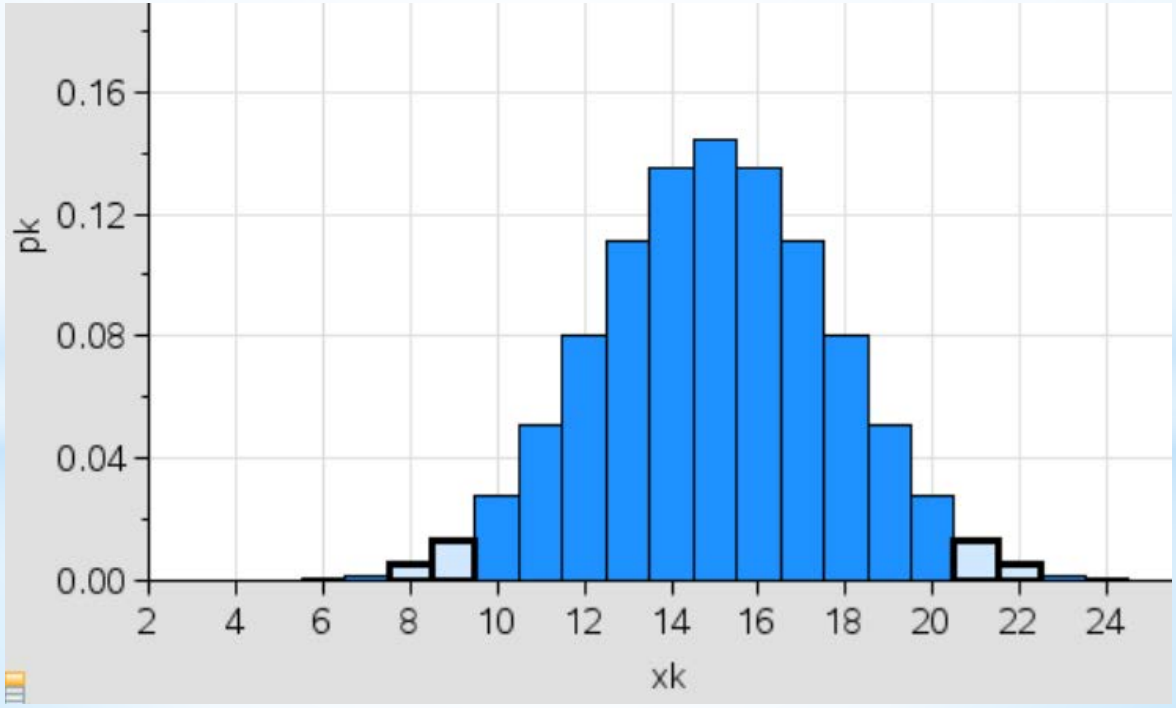
Signifikanter Test: α = 0,05

Hochsignifikanter Test: α = 0,01

3. Nicht-Ablehnungs- und Ablehnungsbereich bestimmen.

Signifikanzniveau α = 5%

binomCdf(30,0.5,0,10)	0.049369
binomCdf(30,0.5,0,9)	0.021387
binomCdf(30,0.5,20,30)	0.049369
binomCdf(30,0.5,21,30)	0.021387



- Die wesentliche Elemente des Test (Signifikanzniveau, Testgröße, kritische Region) sind unabhängig von den Daten, also vor der Auswertung zu bestimmen.
- Da nur die Fehlerwahrscheinlichkeit 1. Art kontrolliert werden kann, kann H_0 nicht mit einer a priori kontrollierten Fehlerwahrscheinlichkeit angenommen, sondern nur abgelehnt, werden.
- Als „guter Forscher“ sollte man deshalb immer das, was man zeigen will, in die Alternativhypothese schreiben.

  z.B. Forscher will zeigen, dass sein Medikament besser wirkt als ein anderes.

  Nullhypothese: Es wirkt schlechter oder gleich gut.
  Alternativhypothese: Es wirkt besser.

  Durch die Kontrolle des Fehlers 1. Art ist gewährleistet, dass es sich der Forscher „möglichst schwer macht“, die Überlegenheit des Medikaments zu zeigen – und umso überzeugender ist es, wenn er es trotzdem schafft.
- Allerdings gibt es nicht immer einen (einfachen) statistischen Test für jede Nullhypothese.

  z.B. ist es technisch viel einfacher als Nullhypothese $\theta = \theta_0$ zu verwenden, als $\theta \neq \theta_0$ ($\rightarrow$ „Äquivalenztest“). Verwendet man deshalb einen Test mit $H_0 : \theta = \theta_0$, möchte man inhaltlich aber genau dies zeigen, kehren sich die Rollen des Fehlers 1. Art und des Fehlers 2. Art um. Um in diesem Fall einen geringeren Fehler 2. Art zu erzielen, sollte das Signifikanzniveau höher als üblich gewählt werden.

Dr. Gero Walter LMU München
http://www.statistik.lmu.de/~walter/lehre/Stat2Soz_09/material/Stat2Soz_09_Skript-Teil3.pdf

Warum sollte man nicht von Annahmebereich reden, sondern den Begriff Nicht-Ablehnungsbereich verwenden?

Variante 2: Nichtablehnungsbereich

{10, 11,…, 19, 20}

Was würde passieren, wenn Ho: p = 0,45 wäre?

binomCdf(30,0.45,0,8)	0.031206
binomCdf(30,0.45,0,7)	0.012105
binomCdf(30,0.45,19,30)	0.033444

Ein weiterer möglicher Zugang



Zwei Kartenspiele werden gemischt und danach wieder in zwei gleich große Spiele geteilt.
Die Nullhypothese „Der Anteil der roten und schwarzen Karten ist gleich“ soll getestet werden.
Entwickeln Sie ein mögliches Testverfahren.