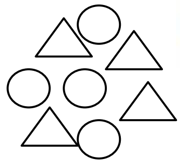
TESTEN VON HYPOTHESEN

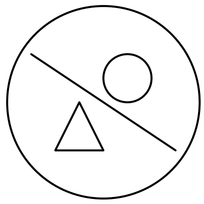
1. Beispiel:
Kann ein neugeborenes Küken Körner erkennen oder lernt es dies erst durch Erfahrung?

Um diese Frage zu entscheiden, wird folgendes Experiment geplant:
Sobald das Küken aus dem Ei schlüpft,
werden ihm falsche Körner aus Papier vorgesetzt.
Die Hälfte sind kleine Kreise, die andere Hälfte
Dreiecke von gleicher Fläche.

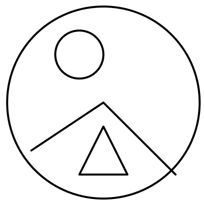


Vor dem Experiment ist das Küken ein unbekannter biologischer Zufallsmechanismus.
Welches Glücksrad steuert sein verhalten?

Man spielt zwei Hypothesen gegeneinander aus:



- H_0 (**die Nullhypothese**):
  Das Verhalten des Kükens wird durch
  ein Laplace-Rad gesteuert.
  Es hat noch nicht gelernt,
  dass Körner rund sind, d.h. $\mathbf{H_0: p = 0{,}5}$



- H_1 (**die Gegenhypothese**):
  Das Verhalten des Kükens wird durch
  das nebenstehende Glücksrad gesteuert.
  Es bevorzugt runde Objakte, obwohl es
  noch nie Körner gekostet hat, d.h. $\mathbf{H_1: p > 0{,}5}$

(Die Möglichkeit $p<0{,}5$ zieht man nicht in Betracht, da sie nicht plausibel erscheint.)

Es soll aufgrund einer Beobachtung zwischen H_0 und H_1 entschieden werden.
Eine Fehlentscheidung lässt sich dabei nicht mit Sicherheit vermeiden.
Jede Beobachtung kann unter beiden Hypothesen stattfinden.

Die Nullhypothese behauptet in der Regel die **Abwesenheit eines Effekts,** den man eigentlich nachweisen will.

Unter einem **Test** versteht man eine Vorschrift, die angibt, ob man sich aufgrund eines Zufallsversuchs für oder gegen H_0 entscheiden soll.

Bei diesem Experiment kann man folgendermaßen vorgehen:
Man lässt das Küken 16mal picken.
Die **Zufallsvariable X** gibt an, wie oft es auf einen Kreis pickt.

Große Werte von X werfen einen Verdacht auf H_0. Was sind nun große Werte für X?

Man kann folgende **Entscheidungsregel** verwenden:

Verwerfe H_0, falls $X \geq 12$
Verwerfe H_1, falls $X > 12$

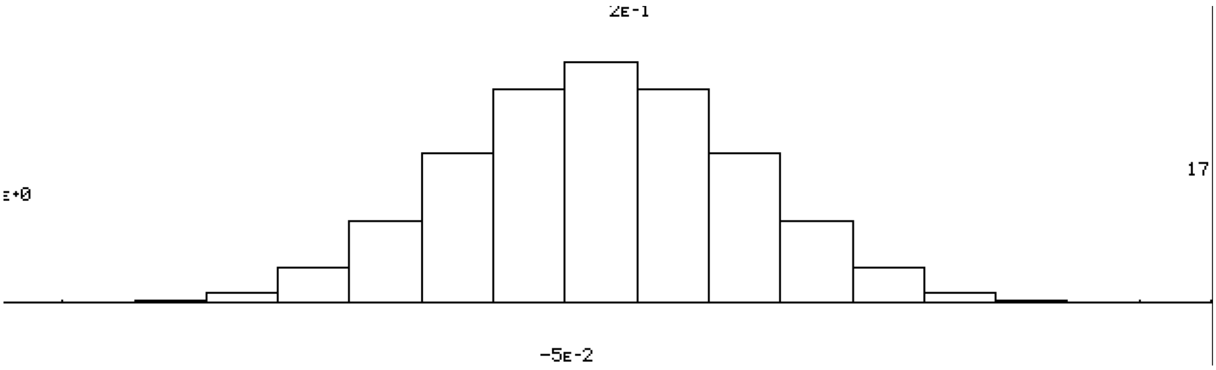
$X \in M = \{$ $\}$
$M = K \cup \overline{K}$,
wobei $K = \{$ $\}$ der Bereich ist, in dem man sich für H_0 entscheidet
und $\overline{K} = \{$ $\}$ der Bereich ist, in dem man H_0 ablehnt.

Nun berechnet man die Wahrscheinlichkeiten der beiden Bereiche:

X ist bei wahrer Nullhypothese $B_{16;0,5}$-verteilt.

$P(X \in K) = P(X < 12) =$

$P(X \in \overline{K}) = P(X \geq 12) =$



Mit dieser Entscheidungsregel sind zwei **Fehlerarten** möglich:

Fehler erster Art: H_0 ist wahr und wird verworfen.

Fehler zweiter Art: H_0 ist falsch und wird angenommen.

Die Wahrscheinlichkeit α, einen Fehler erster Art zu begehen, nennt man **Irrtumswahrscheinlichkeit.**

Man nennt $1-\alpha$ **die statistische Sicherheit**.

Sehr häufig wird $\overline{K}$ so gewählt, dass $\alpha = 0{,}05 = 5\%$.
Dies bedeutet eine statistische Sicherheit von 95%.

In unserem Beispiel:

Fehler 1. Art:

Voraussetzung: H_0 : p = 0,5 ist wahr

H_0 ist wahr und wird verworfen,

d.h. $X \in \overline{K} = \{12,13,14,15,16\}$

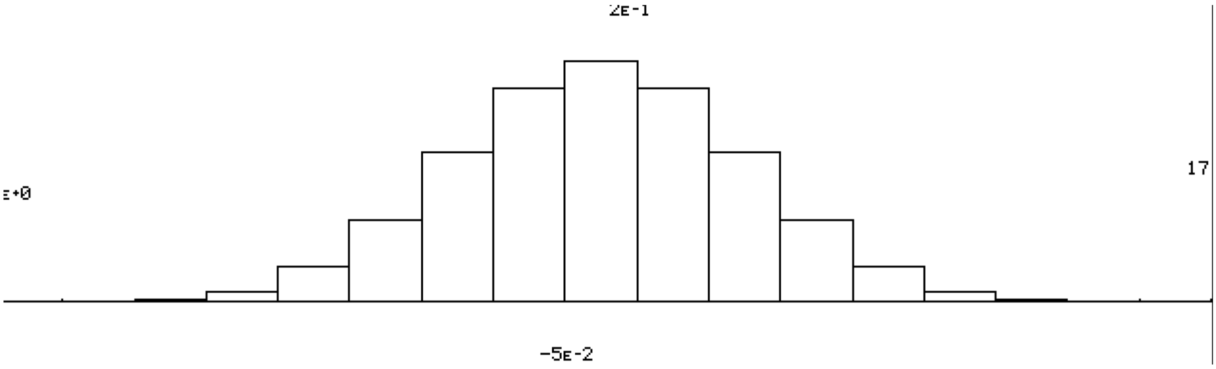
Die **Wahrscheinlichkeit für einen Fehler erster Art** nennt man auch

Risiko erster Art = $P(X \in \overline{K}) = P(X \geq 12) = 0{,}0384 = 3{,}84\% = \alpha$.

Die WS in $\overline{K}$ zu kommen, beträgt 3,84%.

Die Irrtumswahrscheinlichkeit beträgt 3,84%. Diese ist kleiner als 5%.

In knapp 4% aller Fälle wird H_0 zu Unrecht verworfen.



Fehler 2. Art:

Voraussetzung: H_0 ist falsch d.h. **H_1: p > 0,5 ist wahr**
d.h. p = 0,6 oder p=0,7……

H_0 ist falsch und wird angenommen, H_1 ist wahr und wird verworfen,

Die **Wahrscheinlichkeit für einen Fehler zweiter Art** nennt man auch **Risiko zweiter Art =**

Das Risiko zweiter Art muss **für jedes p** berechnet werden:

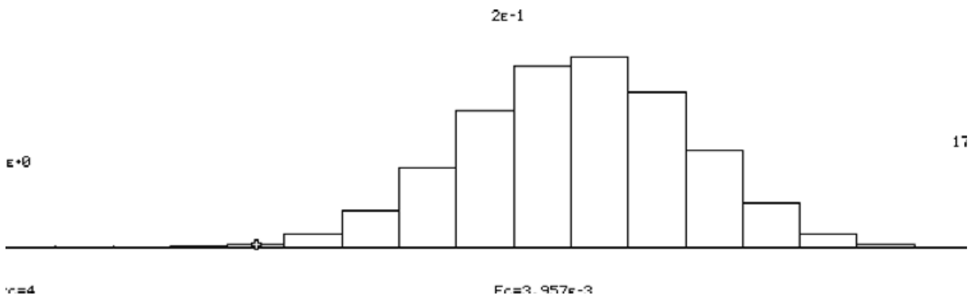
z.B. $p_1 = 0,6$ d.h. X ist $B_{16;0,6}$ - verteilt

$P(X \in K) = P(X<12) =$

$p_1 = 0,7$ d.h. X ist $B_{16;0,7}$ - verteilt

$P(X \in K) = P(X<12) =$

Das Risiko zweiter Art ist abhängig von p.
Beispiel p = 0,7



Allgemeine Bemerkungen zum Testen:

$p =$ unbekannte WS für ein bestimmtes Ereignis E

$H =$ zulässige Hypothese = Menge aller Werte, die p annehmen kann

$H_0 =$ Nullhypothese = ausgezeichnete Teilmenge von H

$H_1 = H \setminus H_0$ Alternativ-Hypothese

$X =$ Zufallsvariable dafür, wie oft E, dessen WS p zu bestimmen ist, bei n Zufallsexperimenten eintritt.

X nimmt Werte aus M = {0,1,2,...,n} an.

$M = K \cup \overline{K}$ mit K: Annahmebereich von H_0

und $\overline{K}$: Ablehnungsbereich von H_0

Entscheidungsregel: Wie wählt man K und $\overline{K}$?

$X \in K$: H_0 wird nicht abgelehnt

$X \in \overline{K}$: H_0 wird abgelehnt

Dabei werden K und $\overline{K}$ festgelegt durch die Vorgabe der Irrtumswahrscheinlichkeit α,

d.h. das Risiko 1.Art soll höchstens gleich α sein.

$P(X \in \overline{K}) \le \alpha$

bzw. $P(X \in K) > \beta = 1 - \alpha$ unter der Voraussetzung, dass H_0 wahr ist

Dabei kann jeder sein Signifikanzniveau wählen:

Entscheidet man sich für das 5% - Niveau und ist $\alpha \le 5\%$, dann muss man H_0 verwerfen. Ist $\alpha > 5\%$, so wird man in der Regel H_0 nicht verwerfen.

Allerdings ist ein α-Wert zwischen 5% und 10% noch „verdächtig“.

Die Annahme oder Ablehnung von H_0 bedeutet nicht, dass H_0 wahr oder falsch ist, sondern nur, dass diese Entscheidung unter den gegebenen Umständen die zweckmäßigste war.

Methodische Tricks und Kniffe - Merkhilfen

Man wählt als Alternativhypothese das, was man vermutet oder bestätigt haben will,
als Nullhypothese das, was abgelehnt werden soll.

Dann wird der Ablehnungsbereich für den Extremfall bestimmt.

Die Art des Ablehnungsbereichs wird der Alternativhypothese entnommen:

H_1: p $>$ ... $\Rightarrow$ rechtsseitiger Test mit Ablehnungsbereich [g;n];
Wertetabelle von $P(X \geq g) = 1 - P(X \leq g-1)$

H_1: p $<$ ... $\Rightarrow$ linksseitiger Test mit Ablehnungsbereich [0;g];
Wertetabelle von $P(X \leq g)$

Und als **ultima ratio**:

Nimm für die Nullhypothese die Stelle im Text, die ein Gleichheitszeichen für eine Wahrscheinlichkeit enthält.