VI Das Testen von Hypothesen

1 Abgrenzung und Problemstellung

Stochastik traditionell Wissenschaft für Spielernaturen. Warum lernt man's dann in der Schule? Anwendung: In der Statistik.

Wahrscheinlichkeitsrechnung: Lehre von den Rechnungen mit Wahrscheinlichkeiten

Statistik: Suche von Modellen aus der Wahrscheinlichkeitsrechnung, die möglichst genau auf real existierende Zufallsprozesse passen.

Typische Aufgabenstellungen der Statistik:
- Wie groß ist die Wahrscheinlichkeit eines konkreten Ereignisses - **Schätzproblem**

Systematische Ermittlung eines Schätzwertes für eine Wahrscheinlichkeit, einen Erwartungswert, eine Varianz bei einem real existierenden Zufallsprozess. Bsp: Wahlhochrechnung
- Stimmt eine angenommene Wahrscheinlichkeit? **– Testproblem**

In beiden Fällen wird der Spezialbegriff **„Stichprobe"** benötigt.
Führt man ein Zufallsexperiment, dessen Ergebnisse mit der Zufallsgröße X beschrieben werden, n mal unabhängig voneinander durch, so spricht man von einer **Stichprobe der Länge n** aus der Zufallsgröße X.

Beispiele:
Test: n- maliges Würfeln, um die Eigenschaften eines Würfels zu testen (Laplace – Würfel oder verdächtig?)
Schätzung: n-maliges Befragen für eine Wahlhochrechnung.

Im Leistungskurs beschäftigen wir uns mit Testproblemen. Einem Test voraus geht die Bildung einer **Hypothese** (Aussage über die Wahrscheinlichkeit eines Ereignisses).

Beispiel:
$P(\{6\}) = 1/6$ beim Wurf eines bestimmten Würfels.
Aufgabe: Eine Entscheidung über die Wahrheit oder Falschheit der Hypothese herbeiführen.
Entscheidungen sind dabei immer fehlerbehaftet.

2 Alternativtest

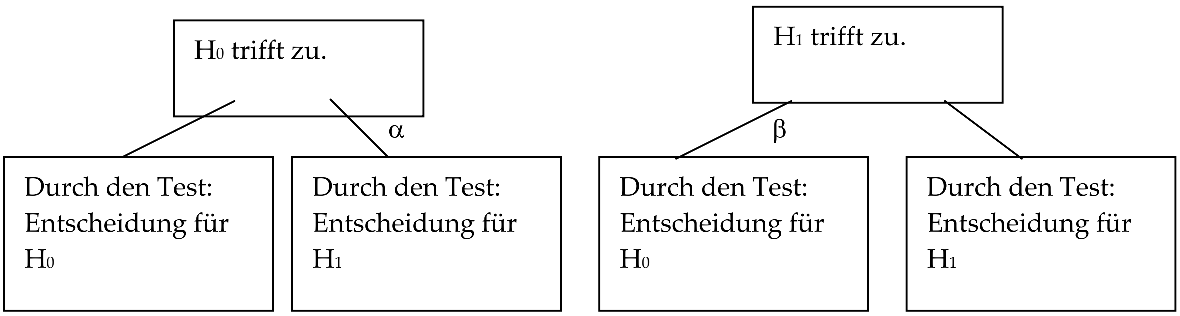
- *Bisher hat eine typische Fragestellung so ausgesehen: 10% einer großen Menge an Schrauben sind fehlerhaft. Wir ziehen 20. Mit welcher W. sind 5 davon kaputt.*
- *Große Anzahl besprechen: Dann egal ob mit oder ohne Zurücklegen.*
- *Nun: Unbekannt wie viel Prozent kaputt sind. Dies soll mit einer Stichprobe ermittelt werden.*
- *Wir vereinfachen in diesem Beispiel das ganze: Zur Auswahl stehen nur zwei mögliche Ausschussanteile: z.B.: 10% und 40%*
- *Daher der Name Alternativtest.*

Beispiel:
Es soll entschieden werden, ob eine große Packung Computerchips dem Qualitätsstandard A (Ausschussanteil: p = 10%) oder dem Standard B (p = 40%) entspricht. Dazu wird eine Stichprobe durchgeführt: 5 Chips werden kontrolliert und man entschließt sich dann für

Hypothese H_0:
„Chips sind von Qualität A (p = 10%)“
Wahrscheinlichkeitsverteilung: B(5; 0,1; k)
falls das Ereignis A: “weniger als 2 sind kaputt“ (=> Annahmebereich {0,1})eintritt; für

Hypothese H_1:
„Chips sind von Qualität B (p = 40%)“
Wahrscheinlichkeitsverteilung: B(5; 0,4; k)
falls das Ereignis $\overline{A}$: “mindestens 2 sind kaputt“ (=> Ablehnungsbereich) eintritt.

Folgende 4 Fälle können eintreten:



α:= Irrtumswahrscheinlichkeit oder Risiko 1. Art
β:= Irrtumswahrscheinlichkeit oder Risiko 2. Art
Es gilt: $\alpha = P^5_{0,1}(\text{"}Z > 1\text{"}) = 0{,}08146$
$\beta = P^5_{0,4}(\text{"}Z \leq 1\text{"}) = 0{,}08146$

Legt man für den Fehler 1. Art eine obere Schranke fest, so nennt man diese **Signifikanzniveau.**

3 Signifikanztest

3.1 Einseitiger Test

Problem:
Entscheidung, ob eine Hypothese H_0 wahr oder falsch ist, ohne dass eine feste Alternative besteht.

Beispiel:
Es soll die Behauptung getestet werden, dass durch Probieren aus einer Teetasse erkannt werden kann, ob zuerst der Tee oder zuerst die Milch in die Tasse gegeben wurde.

- *Vorschläge zu Tests sammeln.*

Test: Aus 10 Tassen wird probiert
Zufallsgröße Z beschreibt, wie viele

Nullhypothese: (*Hypothese, die man nur mit gutem Grund aufgibt)*
„Es gibt keine solche Begabung“; Annahmebereich: „$Z \leq 6$“ {0,1,2,...,6}
Ablehnungsbereich „$Z > 6$“ {7,8,9,10, ...}
Da der Ablehnungsbereich aus einem Intervall besteht, heißt der Test **einseitig**.

Fehler 1 Art:
$P_{0,5}^{10}(Z > 6) = 17\%$
Also: Mit 17% Wahrscheinlichkeit attestieren wir eine solche Begabung zu unrecht.
Diese Wahrscheinlichkeit soll zumeist gering gehalten werden, deshalb wird oft eine obere Schranke α (das Signifikanzniveau) festgelegt.
$1-\alpha$ heißt **statistische Sicherheit** der Nullhypothese.

Fehler 2. Art: Problem: unendlich viele Wahrscheinlichkeitsverteilungen

Beim Signifikanztest ist zunächst nur der Fehler 1. Art interessant und nur das Risiko 1. Art ist berechenbar. (Alternative ist ja nicht bekannt).

3.2 Zweiseitiger Test

Ein Test heißt **zweiseitig**, wenn der Ablehnungsbereich $\overline{A}$ durch den Annahmebereich A in zwei Intervalle geteilt wird.

Beispiel:
Beim Münzwurf lautet die typische Problemstellung, ob es sich um eine L-Münze handelt, ob also Kopf mit einer Wahrscheinlichkeit von 1/2 geworfen wird. Ein sinnvoller Annahmebereich beim zehnmaligen Werfen könnte dann A = {4, 5, 6} sein, also $\overline{A} = \{0, 1, 2, 3\} \cup \{7, 8, 9, 10\}$

Beachte
Das Signifikanzniveau α wird dabei auf beide Bereiche mit α/2 aufgeteilt.

Typ I: Geg: Kritischer Bereich, Ges: α
Urne mit 10 roten und gelben Kugeln.
H_0: genau 5 rote, $\overline{A} = \{0, 1\} \cup \{4, 5\}$
$\alpha = B^5_{0,5}(0) + B^5_{0,5}(1) + B^5_{0,5}(4) + B^5_{0,5}(5) = 36{,}5\%$. Die Nullhypothese wird oft irrtümlich abgelehnt!

Typ II: Geg: α, Ges: Kritischer Bereich
Urne mit 10 roten und gelben Kugeln.
H_0: genau 2 rote, $\alpha = 0{,}2$
$P^5_{0,5}(X \le k) \le \frac{\alpha}{2} = 0{,}1; P^5_{0,5}(X \ge k') \le \frac{\alpha}{2} = 0{,}1 = 36{,}5\%.$
Im Tafelwerk liest man ab: k = 1, k' = 7, also $\overline{A} = \{0, 1\} \cup \{7, 8, 9 ..., 20\}$