ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Statistische Tests

(basierend auf Slides von Lukas Meier)

Statistische Tests: Idee / Ziel

Können wir **basierend auf Daten** entscheiden / nachweisen

- ob ein **Grenzwert** überschritten wird, z.B. bei Asbestfasern?
- ob ein Hersteller bei einem Produkt die **Spezifikationen** verletzt?
- ob sich die gemessene Chlorid-Konzentration an der Oberfläche an einem Betontragwerk von unseren **Modellannahmen** unterscheidet?
- …

Statistische Tests: Idee / Ziel

- Wir wollen also herausfinden, ob eine **bestimmte Annahme** oder ein **bestimmter Parameter** mit unseren **beobachteten Daten** verträglich ist oder nicht.
- Wir müssen eine **objektive, reproduzierbare Entscheidungsregel** verwenden.

Einfachstes Beispiel:

- Sie vermuten, dass eine Münze nicht «fair» ist und **zu oft Kopf** zeigt.
- In 10 neuen Würfen haben wir 8 mal Kopf erhalten. Passt dies mit der Annahme $p = 0.5$ (Münze ist fair) zusammen?

Statistische Tests: Vorgehensweise

- **Modell:** $X =$ Gesamtzahl Köpfe in 10 Würfen. $X \sim Bin(10, p)$
- **Nullhypothese:** $H_0 : p = 0.5$ ($p =$ W'keit für Kopf)
  Dies ist die «Normalzustand» oder «Status Quo».
- **Alternativhypothese:** $H_A : p > 0.5$
  Dies ist was Sie nachweisen wollen, Ihre Vermutung.

Was für Möglichkeiten gibt es?

- H_0 stimmt und es war Zufall, dass wir so oft Kopf gesehen haben.
- H_A stimmt und deshalb haben wir so oft Kopf gesehen.

Statistische Tests: Verschiedene Fehlerarten

Wahrheit \ Entscheidung	H_0	H_A
H_0	✓	Fehler 1. Art (false positive)
H_A	Fehler 2. Art (false negative)	✓

Statistische Tests: Bestimmung Verwerfungsbereich

- Man kann nie sicher sein, welche Entscheidung richtig ist. Aber je öfter man Kopf sieht, je plausibler wird es, dass H_A stimmt.
- Darum macht es sinn H_0 abzulehnen falls $X \geq c$.
- Wie wählt man c?

- Clicker

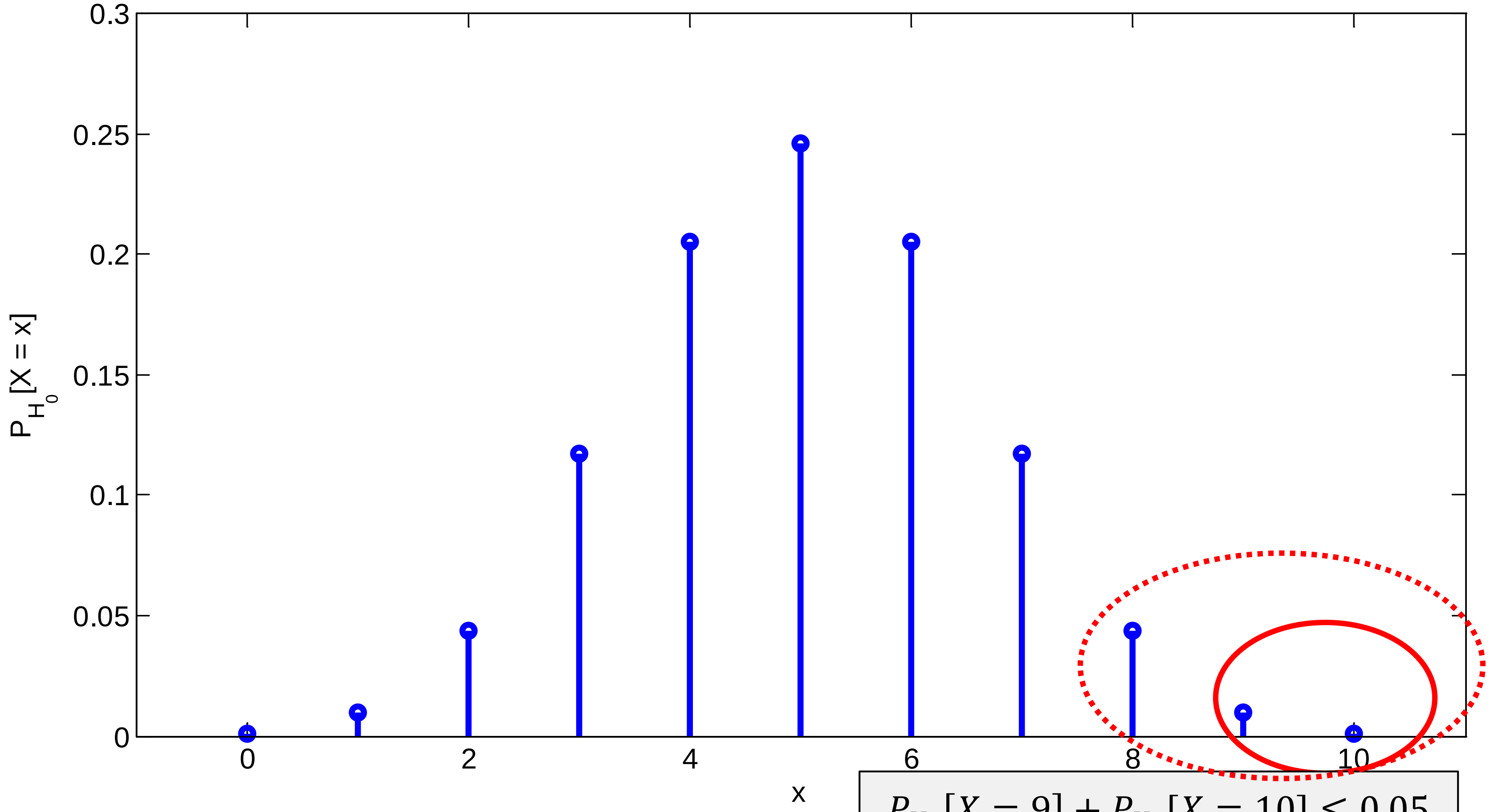
Statistische Tests: Bestimmung Verwerfungsbereich

- Für c klein wird $P(Fehler\ 1. Art) = P(H_0\ verwerfen | H_0\ stimmt)$ gross.
- Für c gross wird es sehr schwierig H_0 abzulehnen, auch wenn H_0 tatsächlich falsch ist. Dadurch wird $P(Fehler\ 2. Art) = P(H_0\ nicht\ verwerfen | H_A\ stimmt)$ gross.
- Kompromiss: Wir wählen das kleinste c_α, so dass
  $$P(Fehler\ 1. Art) = P_{H_0}(X \geq c_\alpha) \leq \alpha$$
  wobei α eine im voraus gesetzte feste Zahl ist, auch “Signifikanzniveau” oder “Niveau” genannt.
- Das signifikanzniveau kontrolliert also $P(Fehler\ 1. Art)$.

- Bemerkung: Man kontrolliert also die Wa’keit für false positives (und nicht die Wa’keit für false negatives)

Statistische Tests: Verteilung unter Nullhypothese

Unter H_0 (d.h. $p = 0.5$) ist $X \sim \mathrm{Bin}(10, 0.5)$ verteilt.



$P_{H_0}[X = 9] + P_{H_0}[X = 10] \leq 0.05$

$P_{H_0}[X = 8] + P_{H_0}[X = 9] + P_{H_0}[X = 10] \geq 0.05$

Statistische Tests: Testentscheid

- Ab $X \geq 9$ verwerfen wir also die Nullhypothese, d.h. der sogenannte **Verwerfungsbereich** ist $\{9,10\}$.
- Wir haben $x = 8$ beobachtet. Das liegt nicht im Verwerfungsbereich. Wir können also die Nullhypothese nicht verwerfen.
- Das heisst jetzt nicht, dass H_0 stimmt oder bewiesen ist. Es heisst nur, dass unsere Daten ($x = 8$) H_0 nicht überzeugend widersprochen haben.
  «Absence of evidence $\neq$ evidence of absence»

Zusammenfassung

- Bestimme Modell (z.B. $X = \#\text{Kopf}$, $X \sim Bin(10, p)$)
- Lege H_0 fest (z.B. $H_0{:}\, p = 0.5$)
- Lege H_A fest (z.B. $H_A{:}\, p > 0.5$)
- Wähle Signifikanzniveau α (z.B. $\alpha = 0.05$)
- Konstruiere Verwerfungsbereich für H_0 (z.B. $\{9{,}10\}$)
- Betrachte, ob die Beobachtung x im Verwerfungsbereich liegt:
  - Ja: verwerfe H_0. Wir glauben an H_A
  - Nein: verwerfe H_0 nicht.

Verwerfungsbereich: Bemerkungen

- Zum Berechnen des Verwerfungsbereichs kann man Grafiken oder Tabellen benutzen. Falls n gross ist, kann man auch die Normalapproximation benutzen.
- Die Form des Verwerfungsbereichs hängt von H_A ab:
  - $H_A: p > p_0 \Rightarrow$ grosse Werte von x unterstützen $H_A \Rightarrow$
    Verwerfungsbereich $x \geq c \Rightarrow$
    wähle den kleinsten c so dass $P_{H_0}(X \geq c) \leq \alpha$.
  - $H_A: p < p_0 \Rightarrow$ kleine Werte von x unterstützen $H_A \Rightarrow$
    Verwerfungsbereich $x \leq c \Rightarrow$
    wähle den grössten c so dass $P_{H_0}(X \leq c) \leq \alpha$.
  - $H_A: p \neq p_0 \Rightarrow$ kleine und grosse Werte von x unterstützen $H_A \Rightarrow$
    Verwerfungsbereich $\{x \leq c_1\} \cup \{x \geq c_2\} \Rightarrow$
    wähle den grössten c_1 und den kleinsten c_2 so dass $P_{H_0}(X \leq c_1) \leq \frac{\alpha}{2}$ und $P_{H_0}(X \geq c_2) \leq \frac{\alpha}{2}$

Statistische Tests: Bemerkungen

- Wir müssen Null- und Alternativhypothese **vor der Datenerhebung** festlegen.
- Sonst könnten wir alle Lotteriegewinner des Betrugs überführen!
- Man kann zwar basierend auf einer Datenanalyse Hypothesen bilden, **um diese zu verifizieren werden aber neue Daten benötigt**.