Statistik

Vorlesung 11 - Hypothesentests

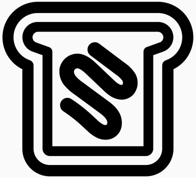
Prof. Dr. Sandra Eisenreich

Hochschule Landshut

Agenda

1. Zweiseitiger Hypothesentest
2. Einseitiger Hypothesentest
3. χ^2-Anpassungstest

Beispiel: Murphy's Law



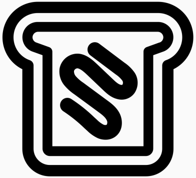
Hypothese: Toast fällt mit $p_0 = 0.5$ auf die Butterseite

n=100 Experimente, Stichprobe: $k = 58$x Butter auf Boden
$\Rightarrow$ geschätzer Parameter $p = 0.58$

Reicht das, um die Hypothese oben abzulehnen?

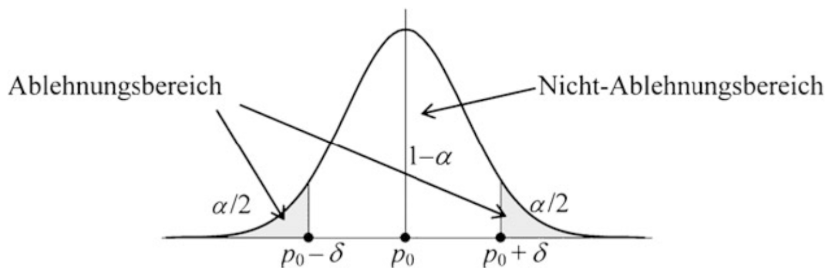
Frage: Ab welcher Stichproben-Abweichung darf man die Hypothese (mit geringer Irrtumswahrscheinlichkeit) ablehnen?

Herleitung: Murphy's Law



Angenommen, wir hätten mit der Hypothese recht.
$\Rightarrow$ X=“wie oft Butter auf Boden” ist b_{100,p_0}−verteilt (bzw. näherungsweise $N(50,5)$-verteilt.)

Idee: Wir lehnen ab, wenn die Stichprobe für diese Verteilung sehr unwahrscheinlich ist. (α : Irrtumswahrscheinlichkeit)



Parametertest

Gegeben:

- Irrtumswahrscheinlichkeit α
- Testfunktion T = Schätzfunktion für den Parameter p
- Hypothese H_0 für den Parameter p

Gesucht: ein Ablehnungsbereich (d.h. ein δ abhängig von α)

Dann wird die Stichprobe durchgeführt. Wir erhalten Schätzwert p.

$$\begin{aligned} p \text{ im Ablehnungsbereich} &\rightarrow \text{Hypothese wird abgelehnt} \\ p \text{ nicht im Ablehnungsbereich} &\rightarrow \text{keine Aussage möglich,} \\ &\quad\ \text{Hypothese wird nicht abgelehnt} \end{aligned}$$

Mögliche Fehler

- Fehler erster Art: Hypothese wird abgelehnt obwohl sie richtig ist. Das ist die Irrtumswahrscheinlichkeit α.
- Fehler zweiter Art: Hypothese wird nicht abgelehnt, obwohl sie falsch ist. Der kann sehr groß sein!

Das Testergebnis ist daher von der Form:

- mit Irrtumswahrscheinlichkeit α kann H_0 abgelehnt werden, oder:
- Das Ergebnis steht nicht im Widerspruch zur Hypothese.

Keinesfalls kann im zweiten Fall die Hypothese angenommen werden.

Zweiseitiger Hypothesentest

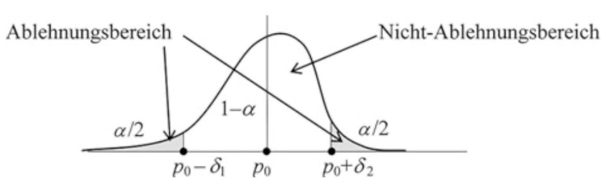
Zweiseitiger Hypothesentest: Allgemeines Vorgehen

Nullhypothese: $H_0 : p = p_0$ (Alternative: $H_1 : p \neq p_0$)

Gegeben: Irrtumswahrscheinlichkeit α, **Gesucht**: δ

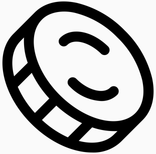
Mit Hilfe der (bekannten) Verteilung werden nun δ_1 und δ_2 berechnet, so dass

$$P(T < p_0 - \delta_1) = \frac{\alpha}{2} \quad \text{und} \quad P(T > p_0 + \delta_2) = \frac{\alpha}{2}$$



Die Verteilung von T muss im Allgemeinen nicht symmetrisch oder eine Normalfunktion sein!

Beispiel: Binomialverteilung, Münze



Teste die Münze eines Spielers:

A= "Münze zeigt Zahl (1)", $p_0 = P(A)$ ist unbekannt.

Hypothese (Nullhypothese) H_0: $p_0 = \frac{1}{2}$

Führe Experiment 100x durch, $X_i \in \{0, 1\}$ i-tes Experiment.

- Die relative Treffer-Häufigkeit $R = \frac{1}{100}(X_1 + \ldots X_{100})$ ist Schätzfunktion für p_0
- Laut Hypothese muss gelten: $E(R) = p_0$ und
- R ist annähernd $\mathcal{N}(p_0, p_0(1 - p_0)/n) = \mathcal{N}(0.5, 0.0025)$-verteilt.

Wir suchen also eine Zahl δ mit $P(p_0 - \delta \leq R \leq p_0 + \delta) = 1 - \alpha$

Allgemeine Berechnung

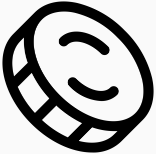
$$\begin{aligned} P(p_0 - \delta \leq R \leq p_0 + \delta) &= \Phi\left(\frac{p_0 + \delta - p_0}{\sqrt{p_0(1-p_0)/n)}}\right) - \Phi\left(\frac{p_0 - \delta - p_0}{\sqrt{p_0(1-p_0)/n)}}\right) \\ &= 2\Phi\left(\frac{\delta}{\sqrt{p_0(1-p_0)/n)}}\right) - 1 = 1 - \alpha \end{aligned}$$

$$\Rightarrow 1 - \frac{\alpha}{2} = \Phi\left(\frac{\delta}{\sqrt{p_0(1-p_0)/n}}\right) \Rightarrow \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = \frac{\delta}{\sqrt{p_0(1-p_0)/n}}$$

Damit erhalten wir schließlich

$$\delta = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\sqrt{p_0(1-p_0)/n}$$

Ergebnis: Binomialverteilung, Münze



Münzwurf; H_0: $p_0 = \frac{1}{2}$
$n = 100$, $\alpha = 10\%$

$$\Rightarrow \delta = \Phi^{-1}\left(1 - \frac{0.1}{2}\right)\sqrt{0.25/100} = 1.65 \cdot 0.05 = 0.083$$

Ablehnungsbereich: $(-\infty, \underbrace{0.5 - 0.083}_{0.417}], \ [\underbrace{0.5 + 0.083}_{0.583}, \infty)$

Jetzt Test durchführen. Bei $k/100 < 0.417$ oder $k/100 > 0.583$ wird Hypothese mit der Irrtumswahrscheinlichkeit von 10% abgelehnt.

Verständnisaufgaben

In obigen Setting, was passiert bei

- bei 41 mal Zahl?
- bei 42 mal Zahl?
- Berechne δ für $n = 200$ und dasselbe α! Was passiert mit dem Ablehnungsbereich?
- Berechne den Ablehnungsbereich für $n = 100$, $\alpha = 1\%$. Wird er größer oder kleiner im Verlgeich zu $n = 100$, $\alpha = 10\%$

Ergebnis: Verständnisaufgaben

- bei 41 mal Zahl $\rightarrow$ Hypothese ist mit Irrtumswahrscheinlichkeit 10% abzulehnen
- bei 42 mal Zahl $\rightarrow$ Ergebnis steht nicht im Widerspruch zur Hypothese!
- $n = 200$ ergibt $\delta = 0.058$, d.h. der Ablehnungsbereich wird größer.
- $n = 100$, $\alpha = 1\%$ ergibt $\delta = 0.128$, d.h. der Ablehnungsbereich wird kleiner (Ablehnung für $k \notin [37, 63]$)

Rechenregel für zweiseitigen Hypothesentest beim Bernoulliexperiment

Rechenregel

In einem Bernoulliexperiment werde die Nullhypothese $P(A) = p_0$ getestet. Die Irrtumswahrscheinlichkeit α sei vorgegeben. Bei n-maliger Ausführung des Experiments trete das Ereignis k-mal ein. Es sei

$$\delta = c \cdot \sqrt{p_0(1 - p_0)/n}, \quad c := \Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$

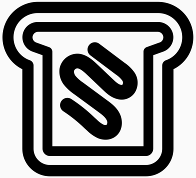
Dann wird für $k/n \notin [p_0 - \delta, p_0 + \delta]$ beziehungsweise für $k \notin [np_0 - n\delta, np_0 + n\delta]$ die Hypothese mit der Irrtumswahrscheinlichkeit von α verworfen.

Wichtige Hinweise

- wenn die Hypothese nicht abgelehnt wird, dann kann sie nicht etwa angenommen werden
- wenn Sie eine Hypothese belegen wollen, sollten Sie den Test so formulieren, dass als Nullhypothese genau das Gegenteil angenommen wird
- Aber: beim zweiseitigen Hypothesentest geht das nicht. $\rightarrow$ einseitiger Hypothesentest

Einseitiger Hypothesentest

Beispiel - Murphy's Gesetz

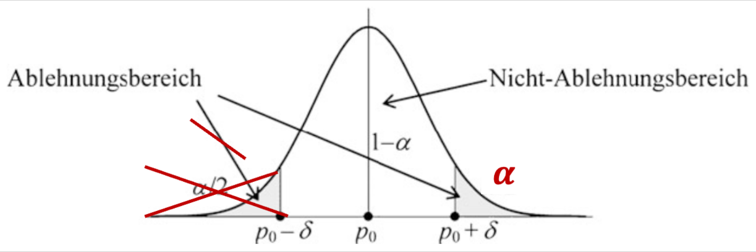


$A =$ "Toast mit Butter auf Boden"

Wollen zeigen: $p = p(A) > \frac{1}{2}$

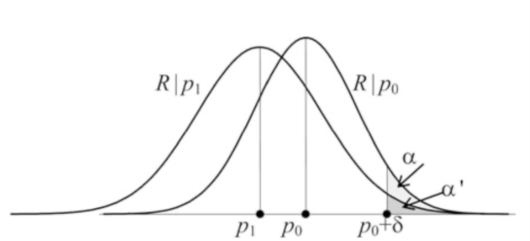
$\Rightarrow$Nullhypothese ist Gegenteil: $H_0 : p \leq \frac{1}{2}$

Nun dürfen wir wegen dem $\leq$ die Hypothese nur noch ablehnen, wenn die Stichprobe auf einer Seite zu weit "außen" liegt!



Beispiel - Murphy's Gesetz

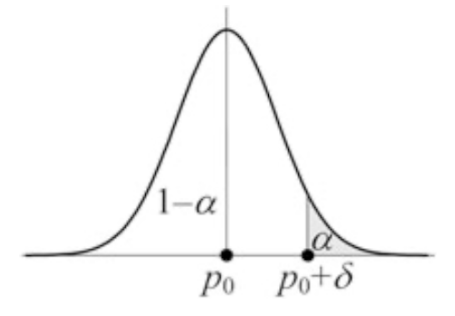
Es reicht, den Grenzfall (hier $\frac{1}{2}$) zu betrachten (wenn eine Stichprobe zu weit rechts ist für die Verteilung mit $p = \frac{1}{2}$, dann erst recht für alle $p < \frac{1}{2}$!)



Einseitiger Hypothesentest:

- es reicht, den Grenzfall zu betrachten,
- berechne hier die maximal zulässige Abweichung δ in eine Richtung.
- Lehne die Hypothese ab, wenn das Ergebnis weiter weg ist.

Die Irrtumswahrscheinlichkeit ist hier nur die "rechte Fläche":



Herleitung Formeln Einseitiger Hypothesentest

Gesucht: δ mit $P(R > p_0 + \delta) = \alpha$:

$$1 - \alpha = P(R \leq p_0 + \delta | p_0) = \Phi\left(\frac{p_0 + \delta - p_0}{\sqrt{p_0(1-p_0)/n}}\right).$$

Wie im zweiseitigen Hypothesentest:

$$\delta = \Phi^{-1}(1-\alpha)\sqrt{\frac{p_0(1-p_0)}{n}} = c \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$$

Rechenregel für einseitigen Hypothesentest beim Bernoulliexperiment

Rechenregel

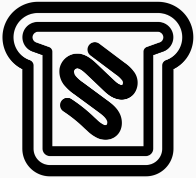
In einem Bernoulliexperiment werde die Nullhypothese $P(A) \leq p_0$ getestet. Die Irrtumswahrscheinlichkeit α sei vorgegeben. Bei n-maliger Ausführung des Experiments trete das Ereignis k-mal ein. Es sei

$$\delta = c \cdot \sqrt{\frac{p_0(1-p_0)}{n}}, \quad c := \Phi^{-1}(1-\alpha).$$

Dann wird für $k/n > p_0 + \delta$ beziehungsweise für $k > np_0 + n\delta$ die Hypothese mit der Irrtumswahrscheinlichkeit von höchstens α verworfen.

Entsprechend wird die Hypothese $p(A) \geq p_0$ für $k/n < p_0 - \delta$ beziehungsweise für $k < np_0 - n\delta$ mit der Irrtumswahrscheinlichkeit von höchstens α verworfen.

Beispiel: Murphy's Gesetz



Wir wollen bestätigen: $P(\text{"Butter auf Boden"}) > \frac{1}{2}$.
Nullhypothese: $H_0 : p \leq \frac{1}{2}$.
Führe Experiment $n = 100$ mal durch. Sei $\alpha = 5\%$.

- Für welche Ergebnisse des Experiments wird die Nullhypothese mit Irrtumswahrscheinlichkeit 5% abgelehnt bzw. bestätigt?
- Wenn die Hypothese abgelehnt wird, können wir dann das Gegenteil mit Irrtumswahrscheinlichkeit 5% als richtig annehmen?
- Wenn die Hypothese nicht abgelehnt werden kann, ist sie dann richtig?
- Hausaufgabe: Wie sieht es aus für $\alpha = 1\%$?

Ergebnis: Murphy's Gesetz

- $\Phi^{-1}(1-0.05) \rightarrow c = 1.64, k = 100, \delta = 1.64 \cdot \sqrt{\frac{1}{2} \cdot \frac{1}{2}/100} = 0.082$
  D.h. für $k > 50 + 8.2 = 58.2$ wird die Hypothese abgelehnt. Beispiele:
  - Resultat 59: Mit Irrtumswahrscheinlichkeit 5% wird H_0 abglehnt, also Murphy hat recht.
  - Resultat 58: Bei Irrtumswahrscheinlicheit 5% steht Resultat nicht im Widerspruch zur Hyptohese.
- Ja
- Nein! Sie kann nur nicht abgelehnt werden.
- $\alpha = 1\% \rightarrow c = 2.33$, $\delta = 0.1165$, Ablehnungbereich $k > 50 + 11.65 = 61.65$

χ^2-Anpassungstest

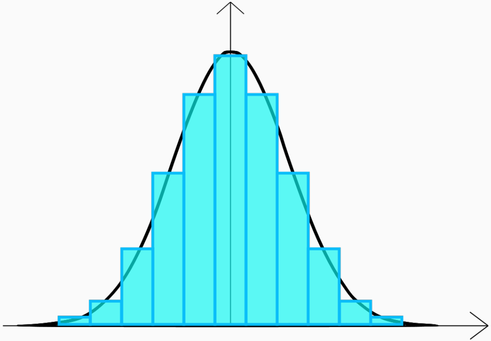
Motivation

- Bei vorliegenden Daten haben wir eine Vermutung, welche Verteilung zugrunde liegt. (häufig: Normalverteilung)
- Wie kann man nachprüfen, ob diese Annahme richtig war?

Dafür gibt es den χ^2-Anpassungstest.

Grundidee

- Teile die Wertemenge der theoretischen Verteilung in disjunkte Abschnitte
- Bestimme deren Wahrscheinlichkeit (Histogramm)
- Damit weiß man, wie oft jeder Abschnitt in einer Stichprobe vorkommen sollte
- Vergleiche mit dem tatsächlichen Ergebnis! Pearsonsche Testfunktion= Maß für die Summe der quadratischen Abweichungen davon



χ^2-Anpassungstest

- x ist Stichprobe einer ZV X, wir vermuten eine Verteilung für X
- Teste diese Vermutung wie folgt:
  - teile den Wertebereich von X in *disjunkte* Teilbereiche $I_1, I_2, \dots, I_m$
  - p_k= Wahrscheinlichkeit für I_k laut Verteilung
  - A_k= "eine Stichprobe liegt in I_k"

Damit lautet die Hypothese H_0, die wir überprüfen wollen:

$$H_0 : P(A_1) = p_1, P(A_2) = p_2, \dots, P(A_m) = p_m$$

Pearson'sche Testfunktion

- X_k sei die Zufallsvariable, die zählt, wie oft bei n Versuchen A_k eintritt
- X_k ist b_{n,p_k}-verteilt $\Rightarrow$ Erwartungswert np_k
- die quadratische Abweichung von X_k vom Erwartungswert im Verhältnis zum Erwartungswert ist also
$$Y_k^2 := \frac{(X_k - np_k)^2}{np_k}$$
- Pearson'sche Testfunktion= Testfunktion für die Nullhypothese:
$$\chi^2 := \sum_{k=1}^{m} Y_k^2 = \sum_{k=1}^{m} \frac{(X_k - np_k)^2}{np_k}$$

Theorem

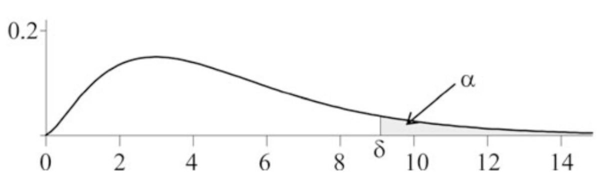
Für $np_k > 5$ ist χ^2 χ^2_{m-1}-verteilt (m= Anzahl von Intervallen).

Pearson's Test - Vorgehen 1. Teil

- Stichprobe von n Experimenten für eine ZV X, wir vermuten eine Verteilung für X
- teile den Wertebereich in *disjunkte* Teilbereiche $I_1, I_2, \ldots, I_m$
- Berechne p_k= Wahrscheinlichkeit für I_k laut Verteilung
- $x_k =$ Häufigkeit der Stichprobe in Intervall I_k
- Berechne $\chi^2 = \sum_{k=1}^{m} \frac{(x_k - np_k)^2}{np_k}$ durch Einsetzen von n, p_k, x_k. Große Abweichungen ergeben große Werte.
- Prüfe ob $np_k > 5$ für alle $k \Rightarrow$ Annäherung durch χ_{m-1} möglich.
- Nun übersetzt sich das Problem in einen einseitigen Hypothesentest, siehe nächste Seite.

Pearson's Test - Vorgehen Hypothesentest

- gebe Irrtumswahrscheinlichkeit α vor und bestimme δ mit $P(\chi^2 > \delta) = \alpha$ durch Ablesen aus der χ^2-Tabelle.



- die Hypothese kann mit der Irrtumswahrscheinlichkeit α abgelehnt werden, wenn bei einer Stichprobe vom Umfang n der berechnete Wert $\chi^2 > \delta$ ist

Beispiel: Würfelwurf

Teste einen Würfel auf Gleichverteilung

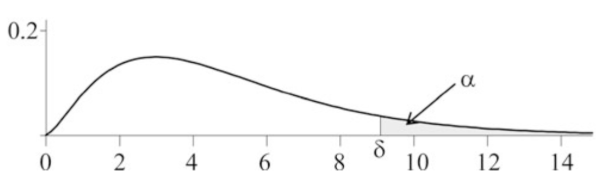
Bei 100-mal würfeln haben wir folgendes Ergebnis erhalten

k	1	2	3	4	5	6
X_k	17	21	17	18	18	9

Kann die Hypothese abgelehnt werden?

Ergebnis: Würfelwurf

- $H_0 : P(A_1) = \frac{1}{6}, P(A_2) = \frac{1}{6}, \ldots, P(A_6) = \frac{1}{6}$
- $\chi^2 = \sum_{k=1}^{6} \frac{(X_k - \frac{100}{6})^2}{\frac{100}{6}} = 4.88$
- Wegen $np_k > 5$ ist unsere Testfunktion annähernd χ_5^2-verteilt.
- Die χ^2-Funktionen sind tabelliert, wir finden darin δ mit $P(\chi_5^2 \leq 9.24) = 0.9$.



Mit einer Irrtumswahrscheinlichkeit von 10% kann die Hypothese nicht abgelehnt werden. Ablehnungsbereich: $(9.24, \infty)$

Verständnisfragen

- Was versteht man unter einer Parameterschätzung?
- Sie wollen ein Konfidenzintervall berechnen. Wenn das Konfidenzniveau angehoben wird, wird dann das Konfidenzintervall kleiner oder größer? Wenn der Umfang der Stichprobe vergrößert wird, wird das Konfidenzintervall kleiner oder größer?

Verständnisfragen

- Was ist der Fehler erster und zweiter Art in einem Hypothesentest?
- Wie sollte man eine Hypothese formulieren, wenn man eine Vermutung bestätigen will?
- Wenn in einem Hypothesentest die Irrtumswahrscheinlichkeit verkleinert wird, wird dann der Ablehnungsbereich kleiner oder größer?

Verständnisfragen - Ergebnisse

- Bei der Verteilung einer Zufallsvariable X ist (mindestens) ein Parameter θ unbekannt. Bei Parameterschätzung schätzen wir diesen Parameter aus einer Stichprobe. Genauer: Die zugrundeliegende Familie möglicher Verteilungen zusammen mit einem Stichprobenraum $\mathcal{X}$ nennt man statistisches Modell $(\mathcal{X}, (P_\theta)_{\theta \in \Theta})$. Ein Schätzer ist eine Abbildung, die einer Stichprobe einen möglichen Wert für *theta* zuweist.
- größere Sicherheit bedeutet kleinere α und größeres Konfidenzintervall. Wird *n* größer, wird das Konfidenzintervall kleiner.

Verständnisfragen - Ergebnisse

- - Fehler erster Art: Hypothese wird abgelehnt obwohl sie richtig ist = Irrtumswahrscheinlichkeit α.
  - Fehler zweiter Art: Hypothese wird nicht abgelehnt, obwohl sie falsch ist.
- Man sollte das Gegenteil als Hypothese formulieren und diese ablehnen.
- Wenn α kleiner wird, dann wird der Ablehnungsbereich kleiner.

Literatur

- Hartmann, Peter; Mathematik für Informatiker, Springer-Vieweg; 7. Auflage; 2019
- Henze, Norbert; Stochastik für Einsteiger; Springer; 10. Auflage; 2013