

UNIVERSITÄT ZU KÖLN

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT



PRAKTIKUM B

B3.1

Statistik der Kernzerfälle

CATHERINE TRAN
CARLO KLEEFISCH
OLIVER FILLA

Inhaltsverzeichnis

1	Einleitung	3
2	Theoretische Grundlagen	4
2.1	Statistik	4
2.1.1	Zufallsvariablen	4
2.1.2	Wahrscheinlichkeitsdichte	4
2.1.3	Wahrscheinlichkeitsverteilung	4
2.2	Wahrscheinlichkeitsverteilungen	5
2.2.1	Binomialverteilung	5
2.2.2	Poissonverteilung	5
2.2.3	Gaußverteilung	6
2.3	Statistische Tests	6
2.3.1	Hypothesentest	6
2.3.2	Fehlerarten	7
2.3.3	Der χ^2 -Anpassungstest	7
2.3.3.1	Pearsons χ^2 -Test	8
2.3.4	Die χ^2 -Verteilung	8
2.4	Versuchs Idee	9
2.4.1	Einfluss der Totzeit	9
3	Durchführung	10
4	Auswertung	11
4.1	Poissonverteilung	11
4.2	Gaußverteilung	11
4.3	Intervallverteilung	11
4.4	χ^2 -Test	11
4.4.1	Hypothese H_1	11
4.4.2	Hypothese H_2	11
4.4.3	Hypothese H_3	11
4.4.4	Halbwertszeit	12
5	Fazit	13
6	Literatur	14

1 Einleitung

In diesem Versuch wird die statistische Methode des χ^2 -Anpassungstests mithilfe von radioaktiver Strahlung untersucht. Dazu wird ^{137}Cs verwendet, dessen Strahlung mit einem Geiger-Müller-Zählrohr detektiert wird.

Damit wird die Hypothese getestet, dass die Präparatstärke im Rahmen der Messung konstant bleibt. Dazu sind die Varianz der gemessenen Werte und die Totzeit des Detektors von essenzieller Bedeutung.

2 Theoretische Grundlagen

2.1 Statistik

2.1.1 Zufallsvariablen

Eine *Zufallsvariable* ist eine Funktion X , die jedem Ereignis ω eines Zufallsexperiments eindeutig eine reelle Zahl zuordnet. Sie kann sowohl diskret als auch kontinuierlich sein.

$$X : \omega \rightarrow x(\omega) \in \mathbb{R} \quad (2.1)$$

Das Ereignis ω des Zufallsexperiments kann direkt eine Zahl sein, wie beispielsweise die Augenzahl eines Würfelswurfs. Alternativ kann Jedem Ereignis des Zufallsexperiments eine Zahl zugeordnet werden, beispielsweise können bei einem Münzwurf dem Ereignis Kopf der Wert 0 und dem Ereignis Zahl eine 1 zugeordnet werden.

2.1.2 Wahrscheinlichkeitsdichte

Eine *Wahrscheinlichkeitsdichte* (PDF¹) ist eine Funktion $f(y)$, welche die Wahrscheinlichkeit \mathbb{P} angibt, dass eine Zufallsvariable X einen Wert innerhalb eines Intervalls $[a, b]$ annimmt.

$$\int_a^b f(y)dy = \mathbb{P}(X(\omega) \in [a, b]) \quad (2.2)$$

Die Wahrscheinlichkeitsdichte kann auch Werte über 1 annehmen. Beispielsweise kann der Ort eines Partikels durch eine δ -Funktion beschrieben werden. Im Fall von diskreten Zufallsvariablen kann auch die Wahrscheinlichkeitsdichte diskret sein.

2.1.3 Wahrscheinlichkeitsverteilung

Eine *Wahrscheinlichkeitsverteilungsfunktion* (CDF²) einer Zufallsvariable X ist dagegen eine Funktion $F(y)$, die angibt mit welcher Wahrscheinlichkeit \mathbb{P} die Zufallsvariable X einen Wert kleiner gleich y annimmt. Sie ist die integrierte Wahrscheinlichkeitsdichte $f(\omega)$.

$$F(y) = \mathbb{P}(X(\omega) \leq y) \quad (2.3)$$

$$F(y) = \int_{-\infty}^y f(\tilde{y})d\tilde{y} \quad (2.4)$$

¹*Probability Density Function*

²*Cumulative Distribution Function*, kumulative Verteilungsfunktion

2.2 Wahrscheinlichkeitsverteilungen

2.2.1 Binomialverteilung

Die *Binomialverteilung* ist eine wichtige diskrete Wahrscheinlichkeitsverteilung. Sie beschreibt ein Zufallsexperiment, das genau zwei sich gegenseitig ausschließende Ereignisse A und B haben kann. Die entsprechenden Wahrscheinlichkeiten \mathbb{P} können wie folgt definiert werden.

$$\mathbb{P}(A) = p \quad (2.5)$$

$$\mathbb{P}(B) = 1 - p \quad (2.6)$$

Dieses Experiment wird N -mal durchgeführt, wobei die einzelnen Ergebnisse jeder Wiederholung unabhängig von den Ergebnissen der vorherigen Wiederholungen sind. Die Zufallsvariable X gibt dann die Anzahl n an eingetretenen Ereignissen A an. Dabei spielt die Reihenfolge, in der die Ereignisse A eintreten, keine Rolle.

Die entsprechende Wahrscheinlichkeit $P(N, n, p)$ gibt demnach die Wahrscheinlichkeit an, dass das Ereignis A genau n -mal eintritt. Weiterhin können der Erwartungswert μ und die Varianz σ^2 bestimmt werden.

$$P(N, n, p) = \binom{N}{n} p^n (1 - p)^{N-n} \quad (2.7)$$

$$m = Np \quad (2.8)$$

$$\sigma^2 = Np(1 - p) \quad (2.9)$$

Der Binomialkoeffizient $\binom{N}{n}$ dient dabei dazu, alle möglichen Reihenfolgen zu berücksichtigen, in der das Ereignis A eintreten kann, wobei der restliche Term die Wahrscheinlichkeit angibt, dass das Ereignis A in einer bestimmten Reihenfolge n -mal eintritt.

2.2.2 Poissonverteilung

Die *Poissonverteilung* beschreibt Reihen von Zufallsvariablen, die unabhängig voneinander eintreten. Sie ist eine weitere diskrete Verteilung mit der Wahrscheinlichkeitsdichte $P(n, \lambda)$, Erwartungswert m und Varianz σ^2 .

$$P_\lambda(n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (2.10)$$

$$m = \lambda \quad (2.11)$$

$$\sigma^2 = \lambda \quad (2.12)$$

Die Poissonverteilung $P_\lambda(n)$ folgt als Grenzfall aus der Binomialverteilung $P(N, n, p)$ mit infinitesimal kleinen Schrittgrößen ($p \rightarrow 0$) und unendlich vielen Schritten ($N \rightarrow \infty$).

Dabei bildet $\lambda \equiv Np$ eine Konstante. Die Poissonverteilung ist als Näherung für die Binomialverteilung zu verwenden, falls folgende Bedingung erfüllt ist.

$$P(n, \lambda) = \lim_{\substack{N \rightarrow \infty \\ p \rightarrow 0}} P(N, n, p) \quad (2.13)$$

$$Np \leq 10 \quad (2.14)$$

2.2.3 Gaußverteilung

Die *Gauß-* ist eine kontinuierliche Verteilung. Ihre Wahrscheinlichkeitsdichte $P(x, \mu, \sigma)$ wird folgendermaßen mithilfe des Erwartungswerts μ und der Varianz σ^2 beschrieben. Die *Normalverteilung* ist eine Gaußverteilung mit $\mu = \sigma^2 = 1$.

$$P(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad (2.15)$$

Der Grenzwert der Poissonverteilung für $\lambda \rightarrow \infty$ liefert die Gaußverteilung.

$$\lim_{\lambda \rightarrow \infty} \frac{\lambda^n e^{-\lambda}}{n!} = \frac{1}{\sqrt{2\pi\lambda}} e^{-\frac{(x-\lambda)^2}{2\lambda}} \quad (2.16)$$

Weiterhin konvergiert die Binomialverteilung nach dem *Satz von Moivre–Laplace* für $N \rightarrow \infty$ mit der Bedingung $0 < p < 1$ gegen die Normalverteilung. Eine Faustregel besagt, dass die Normalverteilung schon eine gute Näherung für die Binomialverteilung liefert, sobald gilt die untenstehende Bedingung erfüllt ist.

$$\lim_{N \rightarrow \infty} \binom{N}{n} p^n (1-p)^{N-n} = \frac{1}{\sqrt{2\pi Np(1-p)}} e^{-\frac{(x-Np)^2}{2Np(1-p)}} \quad (2.17)$$

$$Np(1-p) \geq 9 \quad (2.18)$$

2.3 Statistische Tests

2.3.1 Hypothesentest

Ein Hypothesentest oder Statistischer Test dient dazu, durch eine Hypothese mittels statistischer Messungen zu prüfen.

Dazu verwendet man eine *Nullhypothese*³ H_0 und eine *Gegenhypothese* oder *Alternativhypothese* H_1 , die sich unterscheiden. Ziel des Tests ist es, die Alternativhypothese H_1 zu belegen. Falls dies nicht gelingt, muss man die Nullhypothese H_0 als wahr annehmen. Diese wird nicht überprüft. [2]

Aufgrund der Zufälligkeit der Ereignisse kann es dabei zwei Arten von Fehlern geben. Ein α -Fehler beschreibt das irrtümliche Ablehnen von H_0 , während ein β -Fehler das fälschliche Annehmen von H_0 bezeichnet.

³*Hypothesis to be nullified* [5]

2.3.2 Fehlerarten

Ein *Fehler erster Art* oder α -Fehler beschreibt die fälschliche Ablehnung der Nullhypothese H_0 in einem Statistischen Test. Man nimmt z.B. an, dass ein Würfel gezinkt ist (H_1), obwohl er in Wahrheit fair ist (H_0). Hierbei ist die H_0 die Annahme eines fairen Würfels. Man spricht hier auch von einem *falsch-positiven* Ergebnis. [3]

Ein *Fehler zweiter Art* oder β -Fehler beschreibt umgekehrt die fälschliche Akzeptanz der Nullhypothese H_0 . Beispielsweise geht man davon aus, dass ein Würfel fair ist (H_0), obwohl er tatsächlich unfair ist (H_1). Man spricht hier auch von einem *falsch-negativen* Ergebnis. [3]

Die statistische Signifikanz beschreibt die erlaubte Wahrscheinlichkeit, einen α -Fehler zu begehen. [4] In einem *Alternativtest* dagegen beschreibt die Signifikanz die Wahrscheinlichkeit, einen α - oder einen β -Fehler zu machen. Bei einer Signifikanz Y sind α - und β -Fehler mit einer Wahrscheinlichkeit von je $\frac{Y}{2}$ erlaubt.

2.3.3 Der χ^2 -Anpassungstest

Der χ^2 -Anpassungstest dient dazu, eine Verteilung von Zufallsvariablen A mit einer theoretischen Verteilung zu vergleichen. Man kann mithilfe des Tests bewerten, ob die Zufallsvariablen der Verteilung entsprechen können. Hierbei werden sowohl Fehler 1. Art als auch Fehler 2. Art berücksichtigt.

Die Grundidee dahinter ist, einen Erwartungswert $\langle A \rangle$ und seine Varianz σ_A^2 bewerten zu können. Das Maß für die Abweichung von der Hypothese wird für einen Freiheitsgrad durch χ^2 beschrieben,⁴ was durch die χ^2 -Verteilung beschrieben wird.

$$\chi^2 = \sum_i x_i^2 \quad (2.19)$$

Mithilfe der χ^2 -Verteilung kann eine Signifikanz Y festgelegt werden. Damit kann ein Intervall $[\chi_{\min}^2, \chi_{\max}^2]$ durch die Verteilungsfunktion $F(x, f)$ ermittelt werden. Liegt das ermittelte χ^2 in diesem Intervall, so kann H_1 als signifikant gültig angenommen werden.

$$F(\chi_{\min}^2, f) = 1 - \frac{Y}{2} \quad (2.20)$$

$$F(\chi_{\max}^2, f) = \frac{Y}{2} \quad (2.21)$$

Oft wird die Signifikanz von $Y = 5\%$ gefordert, wodurch das Gültigkeitsintervall durch folgende Gleichungen bestimmt wird.

$$F(\chi_{\min}^2, f) = 0.975 \quad (2.22)$$

$$F(\chi_{\max}^2, f) = 0.025 \quad (2.23)$$

⁴Man könnte auch den Betrag $|x_i|$ anstatt des Quadrates x_i^2 wählen. Dies wird nicht gemacht, weil damit schwieriger zu rechnen ist.

Falls $\chi^2 < \chi_{\min}^2$ das Ergebnis des Tests ist, sind die Daten zu gut an die These angepasst. Dies kann beispielsweise durch Overfitting entstehen.

2.3.3.1 Pearsons χ^2 -Test Eine Variante des χ^2 -Tests betrachtet nur ein Ende der Gauß-Verteilung. Hierbei wird die Signifikanz Y verwendet, um ein maximal gültiges χ_{\max}^2 zu bestimmen, dabei wird auf einen minimalen Wert verzichtet. [?] Damit kann eine Hypothese nur dann abgelehnt werden, wenn χ^2 zu groß ist, ein zu kleines χ^2 ist dabei nicht betrachtet. Auch hier wird oft eine Signifikanz von 5 % verwendet.

$$F(\chi_{\max}^2, f) = Y \quad (2.24)$$

$$F(\chi_{\max}^2, f) = 0.05 \quad (2.25)$$

$$(2.26)$$

2.3.4 Die χ^2 -Verteilung

Sei A standardnormalverteilt⁵, dann ist die χ_1^2 -Verteilung eine quadrierte Normalverteilung mit einem Freiheitsgrad. Daher ist der Erwartungswert $\langle \chi_1^2 \rangle = 1$. Gibt es mehrere Freiheitsgrade f , so müssen f Erwartungswerte $\langle \chi_i^2 \rangle$ addiert werden, um den gesamten Erwartungswert zu ermitteln. Dies wird durch die Wahrscheinlichkeitsdichte (PDF⁶) $f(x, f)$ beschrieben,⁷ wobei die Gammafunktion $\Gamma(x)$ benötigt wird.

$$f(x, 2f) = \begin{cases} \frac{x^{f-1} \exp[-\frac{x}{2}]}{2^f \Gamma(f)} & : x \geq 0 \\ 0 & : x < 0 \end{cases} \quad (2.27)$$

$$\Gamma(x) = \int_0^\infty t^{x-1} \cdot e^{-t} dt \quad (2.28)$$

Die Verteilungsfunktion (CDF⁸) $F(x, f)$ ist dabei komplex und hat den Erwartungswert $\langle \chi_f^2 \rangle = f$ und die Varianz $\sigma_{\chi^2} = 2f$.

$$F(x, 2f) = \int_0^x \frac{y^{f-1} \exp[-\frac{y}{2}]}{2^f \Gamma(f)} dy \quad (2.29)$$

$$\langle \chi_f^2 \rangle = \int_0^\infty x \cdot f(x, f) dx = f \quad (2.30)$$

$$\sigma_{\chi^2} = \int_0^\infty (x - \langle \chi_f^2 \rangle)^2 \cdot f(x, f) dx = 2f \quad (2.31)$$

⁵Diese Annahme ist bei ausreichend vielen Messung durch das Gesetz der großen Zahl gerechtfertigt.

⁶*probability density function*

⁷Achtung: Hier wird zur besseren Lesbarkeit $f(x, 2f)$ angegeben, die Zahl der Freiheitsgrade wird in der Funktion halbiert.

⁸*cumulative distribution function*

2.4 Versuchsidee

In diesem Versuch wird ^{137}Cs als radioaktive Probe verwendet, das eine Halbwertszeit $t_{\frac{1}{2}} \approx 30 \text{ a}$ hat.

Damit soll die folgende Hypothese getestet, die Präparatstärke sei konstant und habe den Wert \bar{n} . Hierbei ist \bar{n} der Mittelwert von vielen Einzelmessungen n_i über eine kurze Zeit von $\Delta t = 20 \text{ s}$, der durch Gleichung (2.32) bestimmt wird. All diese N Messungen werden in einem Zeitraum von wenigen Stunden absolviert.

Da der Zeitraum der Messungen sehr kurz gegen die Halbwertszeit ist, kann man annehmen, dass die Stärke der Probe sich im Rahmen der Messungenauigkeit nicht verändert.

Damit können die Differenzen zum Mittelwert ($n_i - \bar{n}$) ermittelt werden. Nach dem zentralen Grenzwertsatz sind die relativen Differenzen standardnormalverteilt. Dadurch kann die Abweichung χ^2 wie folgt ermittelt werden.

$$\bar{n} = \sum_{i=1}^N \frac{n_i}{N} \quad (2.32)$$

$$\chi^2 = \sum_{i=1}^N \frac{(n_i - \bar{n})^2}{\bar{n}} \quad (2.33)$$

2.4.1 Einfluss der Totzeit

Die Totzeit der Länge τ hat einen Einfluss auf die gemessenen Zählraten. Anstatt einer Zählrate von $\frac{n_i}{\Delta t}$ wird eine totzeitkorrigierte Anzahl k_i gemessen. Dadurch kann ein korrigierter Mittelwert M nach (2.32) bestimmt werden und man erhält eine korrigierte Abweichung χ_{kor}^2 . Die korrigierte Rate wird nach Gleichung (??) bestimmt.

$$k_i = \frac{n_i}{1 - \frac{m}{\Delta t} \tau} \quad (2.34)$$

$$M = \sum_{i=1}^N \frac{k_i}{N} \quad (2.35)$$

$$\chi_{\text{kor}}^2 = \sum_{i=1}^N \frac{(k_i - M)^2}{M} \quad (2.36)$$

Durch Einsetzen der Gleichungen (2.34)-(2.35) sowie (2.33) kann man (2.36) vereinfachen und man erhält die folgende vereinfachte Relation. Hier sieht man, dass die korrigierte Abweichung χ_{kor}^2 kleiner als die nicht-korrigierte Abweichung χ^2 ist, was kontraintuitiv wahrgenommen werden kann.

$$\chi_{\text{kor}}^2 = \frac{1}{1 - \frac{m}{\Delta t} \tau} \cdot \chi^2 \quad (2.37)$$

3 Durchführung

4 Auswertung

4.1 Poissonverteilung

4.2 Gaußverteilung

4.3 Intervallverteilung

4.4 χ^2 -Test

Es werden 51 zufällige Messergebnisse gewählt, aus denen χ^2 gewählt wird. Da der Mittelwert einen statistischen Freiheitsgrad bindet, bleiben 50 Freiheitsgrade übrig, um die Gültigkeitsintervalle zu bilden.

$$\chi_{\min}^2 = 32.357 \quad (4.1)$$

$$\chi_{\max}^2 = 71.420 \quad (4.2)$$

4.4.1 Hypothese H_1

Nach (2.33):

$$\bar{x} = \bar{n} \quad (4.3)$$

$$\chi_1^2 = \sum_i \frac{(x_i - \bar{x})^2}{\bar{x}} \quad (4.4)$$

4.4.2 Hypothese H_2

$$\bar{x}' = 0.9 \cdot \bar{n} \quad (4.5)$$

$$\chi_2^2 = \sum_i \frac{(x_i - \bar{x}')^2}{\bar{x}'} \quad (4.6)$$

4.4.3 Hypothese H_3

Mit $i \in [0, N]$:

$$\langle x(i) \rangle = \bar{n} - i \quad (4.7)$$

$$\chi_3^2 = \sum_i \frac{(x_i - (n - i))^2}{(n - i)} \quad (4.8)$$

4.4.4 Halbwertszeit

$$N(t) = \bar{n} \cdot \exp[-\lambda t] \quad (4.9)$$

$$\bar{n} - 1 = \bar{n} \cdot \exp[-\lambda \Delta t] \quad (4.10)$$

$$\Rightarrow \lambda = \frac{1}{\Delta t} \ln\left(\frac{\bar{n}}{\bar{n} - 1}\right) \quad (4.11)$$

$$\Rightarrow T_{1/2} = \frac{\Delta t \ln(2)}{\ln\left(\frac{\bar{n}}{\bar{n} - 1}\right)} \quad (4.12)$$

5 Fazit

6 Literatur

1. Universität zu Köln, “B3.1: Statistik der Kernzerfälle”, Januar 2021, Online verfügbar unter https://www.ikp.uni-koeln.de/fileadmin/data/praktikum/B3.1_statistik_de.pdf
2. Wikipedia, “Statistischer Test”, https://de.wikipedia.org/wiki/Statistischer_Test, Abruf am 18.04.2024
3. Wikipedia, “Fehler 1. und 2. Art”, https://de.wikipedia.org/wiki/Fehler_1._und_2._Art, Abruf am 18.04.2024
4. Wikipedia, “Statistische Signifikanz”, https://de.wikipedia.org/wiki/Statistische_Signifikanz, Abruf am 18.04.2024
5. G. Gigerenzer, “Mindless statistics”, 2004, *The Journal of Socio-Economics*, p.587-606, DOI 0.1016/j.socec.2004.09.033
6. K. C. Kapur & M. Pecht, “Reliability Engineering”: Appendix E, Wiley 2014, DOI 10.1002/9781118841716
7. E. Cramer & U. Kamps, “Grundlagen der Wahrscheinlichkeitsrechnung und Statistik”, Springer 2020, DOI 10.1007/978-3-662-60552-3
8. J. Puhani, “Statistik”, Springer 2020, DOI 10.1007/978-3-658-28955-3
9. Mary L. McHugh, “The Chi-square test of independence”, *Biochemia Medica*, DOI 10.11613/BM.2013.018