| 1 | 2 | 3 | Σ |
| --- | --- | --- | --- |
| 6/10 | 11,5/14 | 4/6 | 21,5 |

**Datenanalyse für Naturwissenschaftler*Innen**
**Statistische Methoden in Theorie und Praxis** — **Uni Freiburg**
Fernandez Lebron, Andres (RZ: af231) — WiSe 2022/23

## Blatt 11 — 26. Januar 2023

## Aufgabe 1 (Teilchenidentifikation)

Durchquert ein geladenes Teilchen ein Gasvolumen, so erzeugt es in dem Medium Ionisation. Die mittlere Menge hängt dabei von der Masse und dem Impuls des Teilchens ab. Daher können durch Ausarbeitung einer Testhypothese, basierend auf der gemessenen Ionisation im Gasvolumen bei bekanntem Impuls, verschiedene Teilchen identifiziert werden.

Betrachten Sie einen Strahl von Teilchen, welcher entweder Pionen oder Elektronen enthält. Man kann nun, als Funktionen der Ionisation $t$, die Wahrscheinlichkeitsdichtefunktion (WDF) $g(t|H_e)$ der Hypothese, dass es sich um ein Elektron handelt, sowie die WDF $g(t|H_\pi)$ der Hypothese, dass es sich um ein Pion handelt, aufstellen. Man wählt nun Elektronen aus, indem gefordert wird, dass $t \leq t_{cut}$:

Nehmen Sie an, dass die beiden WDFs gegeben sind durch eine um $t = 0$ zentrierte Gaußverteilung für die Elektronen und durch eine um $t = 2$ zentrierte Gaußverteilung für die Pionen. Beide Gaußverteilungen haben eine Standardabweichung von Eins. Ein Test zur Elektronselektion wird konstruiert, indem eine Ionisation $t \leq 1$ gefordert wird.

(a) Wie hoch ist die Signifikanz des Tests? Wie groß ist die Wahrscheinlichkeit, Elektronen zu akzeptieren?

(b) Wie groß ist die Mächtigkeit des Tests gegen die Hypothese, dass das Teilchen ein Pion ist? Wie groß ist die Wahrscheinlichkeit, dass ein Pion als Elektron akzeptiert wird?

(c) Betrachten Sie eine Strahlzusammensetzung von 99% Pionen und 1% Elektronen. Wie groß ist die Reinheit an Elektronen in der durch $t \leq 1$ selektierten Auswahl?

**Lösung**

$$g(t|H_e) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \; ; g(t|H_\pi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-2)^2} \; ; t_{cut} = 1$$

(a)

$$\begin{aligned}
\alpha &= \int_1^\infty g(t|H_e)dt = \int_1^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \\
&= \frac{1}{\sqrt{2\pi}} \cdot \int_1^\infty e^{-\frac{1}{2}t^2} dt \\
&= \frac{1}{\sqrt{2\pi}} \left( \int_0^\infty e^{-\frac{1}{2}t^2} dt - \int_0^1 e^{-\frac{1}{2}t^2} dt \right) \\
&= \frac{1}{\sqrt{2\pi}} \frac{1}{2} \sqrt{2\pi} - 0.3413 \\
&= 0.5 - 0.3413 \\
&= 0.1587
\end{aligned}$$

| Wolfram Alpha?

$$
\begin{aligned}
a &= \int_{-\infty}^{1} g(t|H_e)dt \quad = 1-\alpha \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1} e^{-\frac{1}{2}t^2} dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{1}{2}t^2} dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{1} e^{-\frac{1}{2}t^2} dt \\
&= 0.5 + 0.3413 \\
&= 0.8413 \\
&= 84.13\%
\end{aligned}
$$

3,5 / 3,5

✓

wg symmetrie d. WDF um $t_{cut}$

(b)

$$
\begin{aligned}
1-\beta &= 1 - \int_{-\infty}^{1} g(t|H_\pi)dt \quad ✓ \quad = 1 - \int_{1}^{\infty} g(t|e)\,dt = 1 - 0{,}1587 = 0{,}8413 \\
&= 1 - \int_{-\infty}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-2)^2} dt \\
&= 1 - \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-2)^2} dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t-2)^2} dt \right) \\
&= 1 - \left( \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{2} \cdot \sqrt{2\pi} + 0.1359 \right) \\
&= 0.3641 \quad f
\end{aligned}
$$

$\Rightarrow \beta = p(t \leq t_{cut} | \pi)$

$= 0{,}1587 = \alpha$

2/3,5

Die Wahrscheinlichkeit, dass ein Pion als Elektron akzeptiert wird $= \beta = 1 - (1-\beta) = 0.6359 = 63.59\%$ f

(c)

$\beta \cdot 0.99 \approx 0.6295$ der Pionen werden als Elektronen akzeptiert.

$\alpha \cdot 0.01 \approx 0.0084$ der Elektronen werden akzeptiert.

$\Rightarrow \frac{0.084}{0.6295+0.0084} \approx 0.013$ der Akzeptierten werden Elektronen sein.

f 0,5 / 3

$$
p(e | t \leq t_{cut}) = \frac{p(t \leq t_{cut} | e)\, p(e)}{p(t \leq t_{cut} | e)\, p(e) + p(t \leq t_{cut} | \pi)\, p(\pi)} = 5{,}08\%
$$

11,5/14

## Aufgabe 2 (Zerfallszeit eines Teilchens unter Verwendung des Likelihoodverhältnis)

In einem Experiment wird ein Satz von Zerfallszeiten $\vec{t} = (t_1, t_2, \ldots, t_n)$ eines Teilchens aufgenommen. Betrachten Sie einen Test der Nullhypothese $H_0 : \tau = 1$, wobei $\tau$ die wahre Lebensdauer des Teilchens ist, gegenüber der Alternativhypothese $H_1$: $\tau > 1$.

(a) Wie lauten die Likelihoodfunktion und der Schätzer der Lebensdauer $\hat{\tau}$ für die $n$ Messungen des exponentiellen Zerfalls? Den Schätzer können Sie z.B. über die Maximum-Likelihood Methode bestimmen.

(b) Wie lauten somit die Maxima der zwei Likelihoodfunktionen $L(\tau = 1, \vec{t})$ für $H_0$ und $L(\tau \geq 1, \vec{t})$ für den vollen Parameterraum $H_0 + H_1$?

(c) Zeigen Sie, dass das Likelihoodverhältnis am jeweiligen Maximum $L_{max}$ gegeben ist durch

$$T = \frac{L_{max}(\tau = 1, \vec{t})}{L_{max}(\tau \geq 1, \vec{t})} = \bar{t}^n \exp\left(-n(\bar{t} - 1)\right),$$

wobei $\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i$ der arithmetische Mittelwert der gemessenen Zerfallszeiten ist.

(d) Für große $n$ ist $\bar{t}$ gemäß der Gauß'schen WDF $f_G(\bar{t}; \mu = 1, \sigma^2 = \frac{1}{n})$ verteilt. Berechnen Sie damit den kritischen Wert von $\bar{t}$ für eine Signifikanz von $\alpha = 0.05$. Sie können die Quantiltabelle aus Vorlesung 17 verwenden.

(e) Berechnen Sie den kritischen Wert $T_{\text{krit.}}$ des Likelihoodverhältnisses.

**Lösung**

(a)

3,5/4

$$f(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}}$$

$$L = \prod_{i=1}^{n} f(t_i) = \prod_{i=1}^{n} \frac{1}{\tau} e^{-t_i/\tau} = \frac{1}{\tau^n} e^{-\frac{1}{\tau}\sum_{i=1}^{n} t_i}$$

$$\ln L = \ln\left(\frac{1}{\tau^n} e^{-\frac{1}{\tau}\sum_{i=1}^{n} t_i}\right) = n \ln\frac{1}{\tau} - \frac{1}{\tau}\sum_{i=1}^{n} t_i \quad ✓$$

$$\frac{d \ln L}{d\tau} = \frac{d}{d\tau}\left(n \ln\frac{1}{\tau} - \frac{1}{\tau}\sum_{i=1}^{n} t_i\right) = -\frac{n}{\tau} + \frac{1}{\tau^2}\sum_{i=1}^{n} t_i$$

$$\Rightarrow \frac{1}{\tau^2}\sum_{i=1}^{n} t_i - \frac{n}{\tau}\Big|_{\tau=\hat{\tau}} \overset{!}{=} 0 \quad ✓$$

$$\Rightarrow \hat{\tau} = \frac{1}{n}\sum_{i=1}^{n} t_i \quad ✓$$

(b)

da $\tau = \hat{\tau}$ beim Maximum

1/2

$$H_0 : max(L) = e^{-\sum_{i=1}^{n} t_i} = e^{-n\hat{\tau}}$$

$$H_0 + H_1 : max(L) = \frac{1}{\left(\frac{1}{n}\sum_{i=1}^{n} t_i\right)^{2}} \; ? \; e^{-\frac{1}{n}\sum_{i=1}^{n} t_i \sum_{i=1}^{n} t_i^{-1}} = \frac{1}{\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} t_i\right)}_{\hat{\tau}}{}^{n}} e^{-n}$$

(c)

$$
\begin{aligned}
T &= \frac{L_{max}(\tau = 1, \vec{t})}{L_{max}(\tau \geq 1, \vec{t})} \\
&= \frac{e^{-\sum_{i=1}^{n} t_i}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} t_i\right)^{-n} e^{-n}} \\
&= e^{\overbrace{-\sum\limits_{i=1}^{n} t_i}^{-n\cdot\bar{t}}} e^{n} \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} t_i\right)^{-n}}_{=\bar{t}^{n}} \\
&= \bar{t}^{n} \exp\left(n - n\bar{t}\right) = \bar{t}^{n} \exp\left(-n(\bar{t}-1)\right)
\end{aligned}
$$

das ist sehr unleserlich

$= e^{-n\bar{t}} \cdot e^{n} \cdot \bar{t}^{n}$ ?

2/2

$\blacksquare$

(d)

$\alpha = 0.05$ $1 - 0.05 = 0.95$ $95\%$

Tabelle: $Z_0 = 1.645$

$t_0 = \mu + \sigma Z_0 = 1 + \frac{1}{\sqrt{n}\cdot 1.645}$

$\Rightarrow \bar{t}_{krit} = \frac{\sqrt{n}+1.645}{\sqrt{n}}$

4/4

(e)

$T_{krit} = \left(\frac{\sqrt{n}+1.645}{\sqrt{n}}\right)^{n} \exp\left(-n\left(\frac{\sqrt{n}+1.645}{\sqrt{n}}\right)\right)^{n}$

1/2

= weiter vereinfachen

4/6

## Aufgabe 3 (Studentische $t$-Verteilung – Anwendung in der Blasenkammer)

Die Studentische $t$-Verteilung kann dazu benutzt werden, um auf einem Datensatz eine Nullhypothese $H_0$ zu testen.

Gegeben sei eine Stichprobe vom Umfang $n$ aus einer Gaußverteilung $N(\mu, \sigma^2)$. Falls $\sigma$ bekannt ist, ist die Verteilung für

$$t = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} \tag{1}$$

eine Gaußverteilung $N(0, 1)$. Wenn $\sigma^2$ jedoch nicht bekannt ist, dann ist $t$ gegeben durch:

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}} \tag{2}$$

mit der Stichprobenvarianz

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \,. \tag{3}$$

In diesem Fall ist $t$ nach der Studentischen $t$-Verteilung mit $n - 1$ Freiheitsgraden verteilt.

Betrachten Sie als Beispiel die Messung eines monoenergetischen Strahls von Teilchen mit Impuls $P_0 = 24.90$ GeV$/c$. Dieser trifft auf eine Blasenkammer und durch Messung der Krümmung entlang der Teilchenspur wird der inverse Impuls $1/P_i$ bestimmt. Nehmen Sie an, dass $1/P$ für 20 Teilchen durch zwei verschiedene Detektoren $A$ und $B$ mit den Ergebnissen

$$\begin{aligned} 1/P_A &= (40.12 \pm 0.46) \times 10^{-3} (\text{GeV}/c)^{-1} \text{ und} \\ 1/P_B &= (40.25 \pm 0.25) \times 10^{-3} (\text{GeV}/c)^{-1} \end{aligned}$$

gemessen wurde, wobei der gemessene Wert der Mittelwert der 20 Messungen ist und die Unsicherheit die Stichproben-Standardabweichung.

(a) Was sind, unter Hinzunahme von Gleichung (2), die Werte von $t$ für beide Messungen? (2P)

(b) Wie viele Freiheitsgrade hat jede Messung?

(c) Nutzen Sie die in Abbildung 1 zur Verfügung gestellte Tabelle, um die Grenze der kritischen Region mit einer Signifikanz von $\alpha = 0.05$ zu finden. Bedenken Sie hierbei, dass Sie einen beidseitigen Test durchführen. Wieso muss dieser Test auf zwei Seiten durchgeführt werden?

(d) In Bezug auf den inversen Impuls der einfallenden Teilchen: Sind beide Messungen damit konsistent?

**Lösung**

(a)

A: $t = \frac{\overline{x}-\mu}{s\sqrt{n}} = \frac{40.12\times10^{-3} - \frac{1}{24.90}}{0.46\times10^{-3}\cancel{/}\sqrt{20}} \approx -0.0196$ f $= -0,595$ 1,5/2

B: $t = \frac{\overline{x}-\mu}{s\sqrt{n}} = \frac{40.25\times10^{-3} - \frac{1}{24.90}}{0.25\times10^{-3}\cancel{/}\sqrt{20}} \approx 0.0799$ f $= 1,598$

(b)

$n - 1 = 20 - 1 = 19$

Jede Messung hat 19 Freiheitsgrade. ✓ 1/1

1,5/2

(c)

Tabelle ⇒ Confidence (2 tailed) = 95% and $n = 19$ (Freiheitsgrade) ⇒ $t_{krit} \pm 2.093$

Da es zwei Ablehnungsbereiche gibt und es um zwei (wiederholte) Messungen handelt, es muss '2 tailed' sein. ?

→ es muss ein beidseitiger Test durchgeführt werden

(d)

Man kann sagen, dass beide Messungen ziemlich konsistent sind; A ist allerdings ein weinig besser als B. f

die Frage ist ob $t_A$ und $t_B$ konsistent mit $H_0$ sind. 0/1

$t_A = -0,395$ , $t_B = 1,598$

$\Rightarrow \quad -2,093 \leq t_A, t_B \leq 2,093$

$\Rightarrow$ $t_A$ & $t_B$ können mit $H_0$ nicht abgelehnt werden