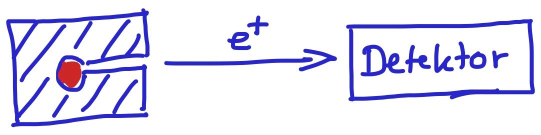
1. Vorlesung, 23. Sept. 2013

Folie "Was ist Datenanalyse?"

Radioaktive Quelle, die eine grosse (aber unbekannte Zahl N von ^{22}Na Atomen enthält.

$^{22}Na \longrightarrow {}^{22}Ne + e^+ + \nu_e$ Halbwertszeit: 2.6027 y



Jedes emittierte Positron gibt 1 "Event" im Detektor. Wir haben innerhalb von 1 Sekunde 20 Events gezählt.

I

Frage: Wie gross ist die Aktivität der Quelle?

Jedes ^{22}Na Atom zerfällt im Zeitintervall Δt mit der Wahrscheinlichkeit $r\Delta t$, wobei r die Zerfallsrate ist. Die Aktivität der Probe entspricht der mittleren Anzahl von Zerfällen pro Sekunde und ist gegeben durch

$A = Nr$. (Einheit: 1 Becquerel = 1 Bq = 1 s^{-1})

Da das Messintervall viel kleiner ist als die Halbwertszeit, ist die Aktivität der Quelle zeitlich praktisch konstant.

Der Zerfall eines einzelnen Atoms geschieht zufällig. Wir fassen das als Zufallsexperiment mit 2 möglichen Ergebnissen auf:

Ergebnis	Wahrscheinlichkeit
Atom zerfällt	$p = r\Delta t$
Atom zerfällt nicht	$1-p = 1-r\Delta t$

Da wir das Zerfallsexperiment mit N Atomen machen, ist das wie beim N-fachen Werfen einer Münze ein Bernoulli-Experiment. Die Wahrscheinlichkeit dabei k Treffer zu erzielen ist durch die Binomial-Verteilung

$$\boxed{\mathrm{prob}(k\,|\,N, r\Delta t, M, I) = B(k; N, r\Delta t) = \binom{N}{k} (r\Delta t)^k (1-r\Delta t)^{N-k} \qquad (1)}$$

M

Folie Binomialverteilung

gegeben. Wir lesen $\mathrm{prob}(k|N, r\Delta t, M, I)$ als "die Wahrscheinlichkeit k Events zu messen, wenn (unter der Bedingung dass) wir N Atome haben (d.h. wir kennen die Zahl N), die Zerfallswahrscheinlichkeit eines Atoms $r\Delta t$ ist (d.h. wir kennen $p = r\Delta t$), dass unser physikalisches Modell M stimmt, und dass unsere Hintergrundinformation I korrekt ist. Wir markieren M und I oben in rot.

Die Wahrscheinlichkeit (1) macht eine Vorhersage über die Zahl der zu erwartenden Events in einem Experiment.

Bevor wir die obige Frage nach der Aktivität der Quelle weiter verfolgen, entwickeln wir noch eine sehr nützliche und gebräuchliche Näherung der Binomialverteilung für grosse N und kleine Δt, unter der Annahme dass $\mu = Nr\Delta t$ endlich bleibt. Wir haben

$$\binom{N}{k}(r\Delta t)^k = \frac{N!}{k!\,(N-k)!}(r\Delta t)^k = \frac{1}{k!}\underbrace{N(N-1)(N-2)\dots(N-k+1)}_{k \text{ Faktoren}}\left(\frac{\mu}{N}\right)^k = \frac{1}{k!}\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\dots\left(1-\frac{k-1}{N}\right)\mu^k \approx \frac{\mu^k}{k!}$$

$$(1-r\Delta t)^{N-k} = \left(1-\frac{\mu}{N}\right)^{N-k} \approx e^{-\mu}$$

$$\Longrightarrow \boxed{\mathrm{prob}(k\,|\,N, r\Delta t, M, I) = P(k;\mu) = \frac{\mu^k}{k!}e^{-\mu} \qquad (2)}$$

Folie Poissonverteilung

Die Verteilungsfunktion $P(k;\mu)$ heisst Poissonverteilung. Wir erinnern uns, dass $\mu = A\Delta t$, wobei A die Aktivität der Quelle und Δt das Messintervall ist.

Gleichung (2) sagt (wie auch (1)) voraus, wie plausibel es ist, k Events zu messen, wenn die Aktivität der Quelle, und auch das Messintervall bekannt sind. Wir stellen aber die umgekehrte Frage: in unseren Fall ist die Zahl der gemessenen Events bekannt, aber die Aktivität ist gesucht. Wie können wir die Aktivität abschätzen?

Eine Analogie:

	US	$\overline{US}$	
G	7.5%	2.5%	10%
$\overline{G}$	60%	30%	90%
	67.5%	32.5%	100%

	US	$\overline{US}$	
G	prob(G,US)	prob(G,$\overline{US}$)	prob(G)
$\overline{G}$	prob($\overline{G}$,US)	prob($\overline{G}$,$\overline{US}$)	prob($\overline{G}$)
	prob(US)	prob($\overline{US}$)	

G: ein Wähler wählt grün

US: ein Wähler befürwortet die Umweltsteuer

$\overline{}$: Negation ($\overline{G}$: ein Wähler wählt nicht grün)

$$\text{prob}(G) = 7.5\% + 2.5\% = 10\%$$
$$\text{prob}(US) = 7.5\% + 60\% = 67.5\%$$

$$\boxed{\begin{aligned}\text{prob}(G) &= \text{prob}(G, US) + \text{prob}(G, \overline{US})\\ \text{prob}(US) &= \text{prob}(G, US) + \text{prob}(\overline{G}, US)\end{aligned}} \quad (3)$$

Marginalisierungsregel der Wahrscheinlichkeitsrechnung

$$\text{prob}(US|G) = \frac{7.5\%}{10\%} = 75\%$$

$$\boxed{\text{prob}(US|G) = \frac{\text{prob}(G,US)}{\text{prob}(G)} \quad (4)}$$ "bedingte Wahrscheinlichkeit"

$$\Rightarrow \text{prob}(G,US) = \text{prob}(G)\,\text{prob}(US|G)$$

$$\text{prob}(G|US) = \frac{7.5\%}{67.5\%} = 11.1\%$$

$$\text{prob}(G|US) = \frac{\text{prob}(G,US)}{\text{prob}(US)}$$

$$\Rightarrow \text{prob}(G,US) = \text{prob}(US)\,\text{prob}(G|US)$$

$$\Rightarrow \boxed{\begin{aligned}\text{prob}(G,US) &= \text{prob}(G)\,\text{prob}(US|G)\\ &= \text{prob}(US)\,\text{prob}(G|US)\end{aligned}} \quad (5)$$

Produktregel der Wahrscheinlichkeitsrechnung

$$\text{prob}(\overline{US}|G) = \frac{2.5\%}{10\%} = 25\%$$

$$\Rightarrow \boxed{\text{prob}(US|G) + \text{prob}(\overline{US}|G) = 1 \quad (6)}$$

Wahrscheinlichkeitsverteilungen integrieren sich zu 1 (= 100%).

wegen (3), (4) und (5) gilt:
$$\text{prob}(US|G) = \frac{\text{prob}(G,US)}{\text{prob}(G)} = \frac{\text{prob}(US)\,\text{prob}(G|US)}{\text{prob}(G,US) + \text{prob}(G,\overline{US})} = \frac{\text{prob}(US)\,\text{prob}(G|US)}{\text{prob}(US)\,\text{prob}(G|US) + \text{prob}(\overline{US})\,\text{prob}(G|\overline{US})}$$

$$\Rightarrow \boxed{\text{prob}(US|G) = \frac{\text{prob}(US)\,\text{prob}(G|US)}{\text{prob}(US)\,\text{prob}(G|US) + \text{prob}(\overline{US})\,\text{prob}(G|\overline{US})}} \quad (7)$$

Bayes'sches Theorem

Das Bayes'sche Theorem erlaubt uns die Invertierung der bedingten Wahrscheinlichkeit prob(G|US) ⟶ prob(US|G). Es spielt eine Schlüsselrolle beim Rückschliessen von Messwerten auf unbekannte Parameter.

Zurück zu unserem Messproblem:

Wir kennen prob(k|A) und können mit Hilfe des Bayes'schen Theorems prob(A|k) bestimmen.

Auf unseren Fall angewendet lautet es:

$$\text{prob}(A|k) = \frac{\text{prob}(A)\,\text{prob}(k|A)}{\sum_A \text{prob}(A)\,\text{prob}(k|A)}$$

Detail: $\text{prob}(A|k) = \underbrace{\text{pdf}(A|k)}_{\text{Wahrscheinlichkeitsdichte}}\,dA$

Der Nenner ist unabhängig von A und sorgt dafür, dass die Wahrscheinlichkeitsverteilung prob(A|k) normiert ist.

Wenn wir uns zunächst einmal auf die Funktionale Abhängigkeit von prob(A|k) beschränken, dann ist

$$\mathrm{prob}(A|k) \propto \mathrm{prob}(A)\,\mathrm{prob}(k|A)$$

Eine Schwierigkeit bei der Anwendung dieser Formel ist, dass wir prob(A) nicht kennen. Was können wir über die Plausibilität verschiedener Werte von A sagen, <u>bevor</u> wir die Zahl k der Events kennen?

Wir wissen, dass $A \geq 0$, denn die Aktivität kann per Definition nie negativ sein. Ansonsten drücken wir die Tatsache, dass wir keine weiteren Anhaltspunkte über den Wert von A haben dadurch aus, dass

$$\mathrm{prob}(A) = \begin{cases} \text{const.} & \text{für } A \geq 0 \\ 0 & \text{für } A < 0 \end{cases}$$

Indifferenzprinzip (Prinzip vom unzureichenden Grund ↗ Laplace 1812)

Nun ist also: $\mathrm{prob}(A|k) \propto \mathrm{prob}(k|A) = \frac{(A\Delta t)^k}{k!} e^{-A\Delta t}$

Wir können diese Verteilungsfunktion für A normieren, indem wir fordern, dass $\sum_A \mathrm{prob}(A|k) = 1$ (siehe Gl. (6))

Wir lösen also: $$\sum_A N \frac{(A\Delta t)^k}{k!} e^{-A\Delta t} = N \int_0^\infty \frac{(A\Delta t)^k}{k!} e^{-A\Delta t}\, dA = \frac{N}{k!\,\Delta t} \int_0^\infty d\mu\, \mu^k e^{-\mu} = \frac{N}{k!\,\Delta t} \times k! = \frac{N}{\Delta t} = 1$$

$$\Rightarrow N = \Delta t$$

und finden daher

$$\boxed{\mathrm{prob}(A|k) = \Delta t \frac{(A\Delta t)^k}{k!} e^{-A\Delta t}\, dA} \quad (8)$$

Folie Wahrscheinlichkeitsdichte für A

Das ist eine Wahrscheinlichkeitsverteilung für die Aktivität A, bei gegebener Zahl von gemessenen Events.
Da die Werte von A beliebige reelle Zahlen ≥ 0 sein können, bezeichnen wir

$$\mathrm{pdf}(A|k) = \Delta t \frac{(A\Delta t)^k}{k!} e^{-A\Delta t}$$ als <u>Wahrscheinlichkeitsdichte</u> von A.

Die Messung erlaubt uns also eine Wahrscheinlichkeitsaussage über die Aktivität A zu machen. Den wahrscheinlichsten Wert für A finden wir als Lösung von

$$\frac{\partial}{\partial A}\mathrm{pdf}(A|k) = 0 \Rightarrow A = \frac{k}{\Delta t} = \frac{20}{1\,\mathrm{s}} = 20\ \mathrm{Bq}.$$

Allerdings hat die Verteilung eine beachtliche Breite um das Maximum herum. Der Wert von A liegt z.B. mit einer Wahrscheinlichkeit von 99.85 % im Intervall [9 Bq; 38 Bq], denn

$$\int_9^{38} \mathrm{pdf}(A|k)\, dA = 0.9985.$$

Folie Zusammenfassung