

TSFA sbírka

Wiki Skriptum FJFI

Martin Štefaňák

11. března 2020

Nejpravděpodobnější rozdělení

Předchozí část nás vybavila znalostmi z teorie pravděpodobnosti a matematické statistiky nezbytných pro vybudování základů statistické fyziky zabývající se statistickými vlastnostmi systému velkého počtu částic. Jako typický názorný příklad můžeme uvést plyn uzavřený v nádobě. Za pokojových podmínek obsahuje 1 mol takového plynu přibližně 10^{23} molekul, pro něž řešení jejich pohybových rovnic je nemyslitelné i s ohledem na to, že nelze stanovit přesné počáteční podmínky takového systému. Měření polohy, rychlosti či případně dalších charakteristik jeho molekul se tak stává náhodným pokusem.

Statistická fyzika, s ohledem na předmět jejího zájmu, zavádí novou terminologii. Výsledek náhodného pokusu, t.j. elementární náhodný jev, se nazývá **mikrostav**. Mikrostavem může tedy být například poloha všech částic, nebo rychlost všech částic, nebo poloha a rychlost všech částic či jejich další charakteristiky (například spin částic). U dříve zmiňovaného hodu kostkou je mikrostavem výsledek hodu. Budeme-li uvažovat náhodný pokus s N opakováními tohoto hodu kostkou, pak mikrostavem takového náhodného pokusu je každá možná N -tice výsledků.

Mikrostav je tedy dán tím, jaký náhodný pokus uvažujeme. My nadále budeme předpokládat, že množina všech možných makrostavů Ω náhodného pokusu je podmnožinou $\mathbb{R}^n$. Náhodný pokus je tedy zároveň vektorem náhodných veličin s příslušnou rozdělovací funkcí jeho možných mikrostavů. Tuto rozdělovací funkci ve statistické fyzice nazýváme **stavem** (nebo také **makrostavem**) statistického systému. Stav systému, t.j. dané pravděpodobnostní rozdělení jeho mikrostavů, pak určuje střední hodnoty všech relevantních náhodných veličin (makroskopické hodnoty pozorovatelných).

Jedním z hlavních úkolů statistické fyziky a ústředním bodem této kapitoly je právě nalézt toto rozdělení mikrostavů odpovídající makroskopickým vlastnostem statistického systému (zadané střední hodnoty náhodných pozorovatelných). Toto rozdělení musí splňovat dvě vlastnosti:

1. Stav systému je v souladu se zadanou informací o systému, tedy v souladu se známými středními hodnotami náhodných veličin.
2. Stav systému maximalizuje neznalost o systému.

První z těchto vlastností jsme již komentovali. Druhá vlastnost je naopak něco, co si zasluhuje podrobněji vysvětlit. Střední hodnoty zadaných náhodných pozorovatelných představují jedinou informaci, kterou o statistickém systému máme. Kdyby stav systému (pravděpodobnostní

rozdělení makrostavů) obsahoval více informace než nezbytné minimum odpovídající vlastnosti 1), znamenalo by to, že o stavu systému toho víme více, například střední hodnoty jiných náhodných veličin či jiný typ informace. My ale znalost další takové informace nepředpokládáme a tedy stav systému musí v jistém smyslu minimalizovat znalost o systému a tedy maximalizovat neznalost o něm. Nezbytným předpokladem k naplnění vlastnosti 2. je, že existuje způsob, jak tuto neznalost kvantifikovat. V další části si takovou míru neznalosti stavu systému zkonstruujeme.

0.1 Míra informace, entropie

Mějme náhodný pokus s množinou mikrostavů Ω . Naším úkolem je přiřadit každému náhodnému jevu $A \subset \Omega$ jeho informační obsah (někdy také nazývaný míra překvapení) vyjadřující **míru informace**, pokud tento jev nastane. Tato míra informace má řadu vlastností, které od ní přirozeně očekáváme. Je zřejmé, že pokud náhodný jev $A \subset \Omega$ je jev jistý, je míra informace takového jevu nulová. Tento jev nastane při každém náhodném pokusu a tedy nenese žádnou informaci, míra informace takového jevu je nulová. S klesající pravděpodobností míra překvapení a tedy množství získané informace roste. Z uvedeného je zřejmé, že tato míra informace daného jevu A je pouze funkcí pravděpodobnosti tohoto jevu, t.j.

$$I(A) = I(P(A)), \quad (1)$$

s následujícími vlastnostmi

1. $I(p) \geq 0$
2. $I(1) = 0$
3. $I(p)$ je klesající funkce.
4. Uvážíme-li dvě **nezávislá** opakování náhodného pokusu, pak míra informace náhodného jevu, že v prvním pokusu nastane jev A a v druhém pokusu jev B , je prostým součtem informací měřených pro každý jev zvlášť, tedy

$$I(A, B) = I(A) + I(B) \quad \Rightarrow \quad I(P(A)P(B)) = I(P(A)) + I(P(B))$$

Tyto čtyři vlastnosti jednoznačně určují tvar diferencovatelné míry informace $I(p)$. Skutečně, zderivujeme-li obě strany rovnice $I(xy) = I(x) + I(y)$ podle proměnné y a dosadíme-li $y = 1$, dostaneme jednoduchou diferenciální rovnici

$$I'(x) = \frac{1}{x} I'(1).$$

Z požadovaných vlastností 1) až 3) plyne,

$$I(p) = -k \ln p,$$

kde k je libovolná kladná konstanta.

Nalezli jsme tedy míru informace náhodného jevu. Míra neznalosti stavu statistického systému pak definujeme jako střední hodnotu míry informace pro tento náhodný pokus. Takové míře neznalosti systému říkáme **entropie**, označuje se S a pro spočetný počet mikrostavů (diskrétní náhodná veličina) má tvar

$$S = \langle I \rangle = -k \sum_{\gamma \in \Omega} p_{\gamma} \ln p_{\gamma}.$$

Je-li některá z pravděpodobností $p_{\gamma} = 0$, pak příspěvek odpovídajícího členu v entropii se z důvodu spojitosti definuje jako $p_{\gamma} \ln p_{\gamma} = 0$. Z matematického pohledu je entropie funkcionál na prostoru pravděpodobnostních rozdělení. Maxima nabývá pro rovnoměrné rozdělení.

Příklad: Balíček 52 karet

Známe-li pořadí karet v balíčku, je entropie jeho rozdělení nulová. Pokud ale karty zamícháme tak, že nevíme, jak jdou za sebou, jsou všechna pořadí karet stejně pravděpodobná, tj. pořadí karet je dáno rovnoměrným rozdělením. Celkový počet možností (mikrostavů) je $|\Omega| = 52!$, takže pro všechny mikrostavy γ je $w_{\gamma} = \frac{1}{52!}$. Entropie rozdělení pořadí karet po zamíchání tedy vzroste na

$$S = -k \sum_{\gamma \in \Omega} \frac{1}{52!} \ln \frac{1}{52!} = k \ln 52!,$$

což je maximální možná hodnota.

Pro systém se spojitou množinou mikrostavů $x \in \mathcal{X} \subset \mathbb{R}$ s hustotou pravděpodobnosti $w(x)$ je entropie definována analogicky vztahem

$$S = - \int_{\mathcal{X}} w(x) \ln w(x) dx.$$

Vybaveni nyní již mírou neznalosti (někdy označovanou také jako míru neuspořádanosti) systému, můžeme se vrátit k našemu původnímu problému. V jakém stavu se nachází systém o němž známe toliko střední hodnoty vybraných náhodných pozorovatelných $A_j, j = 1, \dots, n$ definovaných na mikrostavech tohoto systému. Hledané rozdělení, tzv. **nejpravděpodobnější** rozdělení, musí splňovat zadané střední hodnoty a navíc maximalizovat entropii, t.j. neznalost o systému. Úloha vede na vázaný extrém funkcionálu entropie S , jehož řešení budeme diskutovat zvlášť pro diskrétní a spojitou množinu mikrostavů.

0.2 Diskrétní množina mikrostavů

Máme zadaný systém se spočetným počtem mikrostavů $\gamma \in \Omega$ a známe střední hodnoty $\langle A_j \rangle$ náhodných veličin $A_j, j = 1, \dots, n$ definovaných na těchto mikrostavech, t.j. $A_j : \gamma \rightarrow A_j^{\gamma}$. Hledáme nejpravděpodobnější rozdělení $p_{\gamma}, \gamma \in \Omega$ za vazbových podmínek dané středními hodnotami

$$\sum_{\gamma \in \Omega} A_j^{\gamma} p_{\gamma} = \langle A_j \rangle, \quad j = 1, \dots, n. \quad (2)$$

Rozdělení musí být navíc normováno k jedné, což dává jednu další vazbu

$$\sum_{\gamma \in \Omega} p_{\gamma} = 1. \quad (3)$$

Nejpravděpodobnější rozdělení za podmínek (2),(3) je dáno vázaným extrémem entropie, který určíme pomocí Lagrangeovy funkce

$$\Lambda = -k \sum_{\gamma \in \Omega} p_{\gamma} \ln p_{\gamma} - k\alpha \left(\sum_{\gamma \in \Omega} p_{\gamma} - 1 \right) - k \sum_{j=1}^n \lambda_j \left(\sum_{\gamma \in \Omega} A_j^{\gamma} p_{\gamma} \right).$$

λ_i jsou Lagrangeovy multiplikátory přiřazené k jednotlivým vazbám (2) a α je Lagrangeův multiplikátor přiřazený normovací podmínce (3). Z podmínky na extrém Lagrangeovy funkce $\frac{\partial \Lambda}{\partial p_{\gamma}} = 0$ dostaneme

$$p_{\gamma} = e^{-1-\alpha} \exp \left(- \sum_{j=1}^n \lambda_j A_j^{\gamma} \right).$$

Z normalizační podmínky (3) navíc získáme

$$\sum_{\gamma \in \Omega} p_{\gamma} = e^{-1-\alpha} \sum_{\gamma \in \Omega} \exp \left(- \sum_{j=1}^n \lambda_j A_j^{\gamma} \right) = 1,$$

z čehož plyne

$$Z \equiv e^{1+\alpha} = \sum_{\gamma \in \Omega} \exp \left(- \sum_{j=1}^n \lambda_j A_j^{\gamma} \right).$$

Výraz Z označuje partiční sumu (Zustandsumme). Nejpravděpodobnější rozdělení má tedy tvar

$$p_{\gamma} = \frac{1}{Z} \exp \left(- \sum_{j=1}^n \lambda_j A_j^{\gamma} \right) \quad (4)$$

Lagrangeovy multiplikátory λ_j se určí dosazením p_{γ} do vazbových podmínek (2).

0.3 Spojitá množina mikrostavů

Uvažujme nyní systém s mikrostavy $x \in \mathcal{X} \subset \mathbb{R}$. V tomto případě úloha nalézt nejpravděpodobnější rozdělení znamená nalézt hustotu pravděpodobnosti $w(x)$ splňující vazebné podmínky dané středními hodnotami a normovací podmínkou

$$\langle A_j \rangle = \int_{\mathcal{X}} A_j(x) w(x) dx, \quad j = 1, \dots, n, \quad (5)$$

$$\int_{\mathcal{X}} w(x) dx = 1. \quad (6)$$

S využitím znalostí variačního počtu (Teoretická Fyzika 2) nalezneme vázaný extrém funkcionálu entropie za podmínek (5),(6) přechodem k funkcionálu

$$\Lambda = -k \int_{\mathcal{X}} w(x) \ln w(x) dx - k \sum_j \lambda_j \left(\int_{\mathcal{X}} A_j(x) w(x) dx - \langle A_j \rangle \right) - k\alpha \left(\int_{\mathcal{X}} w(x) dx - 1 \right).$$

Definice Lagrangových multiplikátorů je analogická jako u diskrétního případu. Variace funkcionálu Λ je rovna

$$\delta\Lambda = -k \int_{\mathcal{X}} \left(1 + \ln w(x) + \sum_j \lambda_j A_j(x) + \alpha \right) \delta w dx.$$

Z podmínky na extrém $\delta\Lambda = 0$ dostaneme

$$\ln w(x) = -1 - \alpha - \sum_j \lambda_j A_j(x).$$

Nejpravděpodobnější rozdělení má tedy tvar

$$w(x) = e^{-1-\alpha} \exp \left(- \sum_j \lambda_j A_j(x) \right) = \frac{1}{Z} \exp \left(- \sum_j \lambda_j A_j(x) \right), \quad (7)$$

kde jsme opět zavedli partiční sumu

$$Z \equiv e^{1+\alpha} = \int_{\mathcal{X}} \exp \left(- \sum_j \lambda_j A_j(x) \right) dx.$$

Lagrangeovy multiplikátory λ_j se určí dosazením $w(x)$ do vazbových podmínek (5). Shledáváme, že v obou případech (diskrétní a spojitá množina mikrostavů), má nejpravděpodobnější rozdělení formálně stejný tvar.

Poznamenejme, že řešení hustoty pravděpodobnosti s více proměnnými vede rovněž na formálně stejné řešení, viz probíraný variační počet v TEF2.

0.4 Příklady

Příklad 0.1. Důležitou vlastností entropie je její *aditivita*. Tato vlastnost říká, že entropie systému složeného z konečného počtu nezávislých podsystémů je rovna součtu entropií jednotlivých podsystémů. Dokažte toto tvrzení.

Návod: Uvažujme například systém složený z N podsystémů se spojitou množinou mikrostavů. Nechť pro jednoduchost jsou mikrostavy i -tého podsystému indexovány reálným parametrem x_i . Mikrostav celého systému je pak definován jako úspořádaná N -tice $(x_1, x_2, \dots, x_N)$.

Rozdělení mikrostavů je hustota pravděpodobnosti $w(x_1, x_2, \dots, x_N)$, která, protože jde o nezávislé podsystémy, splňuje

$$w(x_1, x_2, \dots, x_N) = \prod_{i=1}^N w_i(x_i),$$

kde $w_i(x_i)$ je marginální hustota pravděpodobnosti i -tého podsystému. S využitím této vlastnosti je důkaz aditivity přímočarý.

Příklad 0.2. Uvažujme šestistěnou kostku, u které 1 padá dvakrát častěji než 6. Najděte nejpravděpodobnější rozdělení výsledků hodu kostkou.

Návod: Vazbové podmínky mají v tomto případě tvar

$$p_1 = 2p_6, \quad \sum_{i=1}^6 p_i = 1.$$

Řešení vázaného extrému entropie hledáme opět v podobě obyčejného extrému Lagrangeovy funkce s příslušnými Lagrangeovými multiplikátory. Nicméně, ač se to možná na první pohled nezdá, lze první vazbovou podmínku interpretovat jako střední hodnotu jisté náhodné pozorovatelné, pro kterou již známe obecné řešení. Skutečně, přepsáním vazebné podmínky do tvaru $p_1 - 2p_6 = 0$ se nabízí nadefinovat náhodnou veličinu A vztahy

$$A^1 = 1, \quad A^2 = -2, \quad A^3 = A^4 = A^5 = A^6 = 0.$$

Pozor, nejde v žádném případě o mocniny, jde o hodnoty náhodné veličiny A pro mikrostavy $i \in \hat{6}$. Vazbová podmínka $p_1 - 2p_6 = 0$ je pak ekvivalentní požadavku $\langle A \rangle = 0$. S využitím již známého obecného řešení (4) zjišťujeme, že pravděpodobnostní rozdělení výsledků hodu má tvar

$$\begin{aligned} p_i &= e^{-1-\alpha} = \frac{1}{Z}, \quad i = 2, 3, 4, 5 \\ p_1 &= \frac{1}{Z}e^{-\lambda}, \quad p_6 = \frac{1}{Z}e^{2\lambda}. \end{aligned}$$

Po dosazení do vazbových podmínek navíc dostáváme, že Lagrangeův multiplikátor λ a partiční suma Z jsou rovny

$$\lambda = -\frac{1}{3} \ln 2, \quad Z = 4 + 2^{\frac{1}{3}} + 2^{-\frac{2}{3}}.$$

Příklad 0.3. Mějme částici na ose x . Víme, že její střední hodnota polohy je rovna μ a střední kvadratická odchylka polohy je σ . Určete nejpravděpodobnější rozdělení polohy částice.

Návod: Hledáme nejpravděpodobnější rozdělení $w(x)$, $x \in \mathbb{R}$ za podmínek

$$\langle x \rangle = \mu, \quad \langle x^2 \rangle = \sigma^2 + \mu^2, \quad \int_{\mathbb{R}} w(x) dx = 1.$$

Nejpravděpodobnější rozdělení má tvar (viz. (7))

$$w(x) = \frac{1}{Z} e^{-\lambda_1 x - \lambda_2 x^2} = \frac{1}{Z} \exp \left[-\lambda_2 \left(x + \frac{\lambda_1}{2\lambda_2} \right)^2 \right] e^{\frac{\lambda_1^2}{4\lambda_2}}.$$

Partiční sumu a Lagrangeovy multiplikátory λ_j získáme dosazením $w(x)$ do vazbových podmínek. Postupně nalezneme

$$\begin{aligned} \int_{\mathbb{R}} w(x) dx = 1 &\implies Z = e^{\frac{\lambda_1^2}{4\lambda_2}} \sqrt{\frac{\pi}{\lambda_2}}, \\ \int_{\mathbb{R}} x w(x) dx = \mu &\implies -\frac{\lambda_1}{2\lambda_2} = \mu, \\ \int_{\mathbb{R}} x^2 w(x) dx = \sigma^2 + \mu^2 &\implies \sigma^2 = \frac{1}{2\lambda_2}. \end{aligned}$$

Lagrangeovy multiplikátory jsou tedy rovny

$$\lambda_1 = -\frac{\mu}{\sigma^2}, \quad \lambda_2 = \frac{1}{2\sigma^2}.$$

Po dosazení se nejpravděpodobnější rozdělení zjednoduší na tvar

$$w(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x - \mu)^2}{2\sigma^2} \right),$$

což je Gaussovo normální rozdělení s parametry μ a σ .

Příklad 0.4. Mějme jednoatomový ideální plyn v nádobě, která je v klidu. Plyn má teplotu T a nenachází se v žádném vnějším poli. Určete nejpravděpodobnější rozdělení rychlostí atomů plynu.

Návod: Mikrostav plynu složeného z N jednoatomových molekul je charakterizován uspořádanou N -ticí vektorů rychlosti $\vec{V} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_N)$, kde $\vec{v}_i = (v_{i1}, v_{i2}, v_{i3})$ je vektor rychlosti i -tého atomu. Rozdělení mikrostavů celého systému je tedy dáno hustotou pravděpodobností $W(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_N)$. V našem případě jde ale navíc o ideální plyn, kdy jednotlivé atomy spolu neinteragují a výsledek měření vektorů rychlosti jednotlivých atomů jsou tedy navzájem nezávislé jevy. Hustotu pravděpodobnosti lze tedy psát ve tvaru

$$W(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_N) = \prod_{i=1}^N w(\vec{v}_i),$$

kde $w(\vec{v}_i)$ je marginální hustota pravděpodobnosti vektoru rychlosti i -tého atomu. Všechny atomy jsou stejné a stejné jsou tedy i funkce charakterizující jejich marginální hustoty pravděpodobnosti.

Protože ale i různé kartézské složky vektoru rychlosti jednoho atomu jsou nezávislé náhodné veličiny a žádný směr není nijak preferovaný, bude platit pro hustotu pravděpodobnosti vektoru rychlosti i -tého atomu

$$w(\vec{v}_i) = w_1(v_{i1})w_1(v_{i2})w_1(v_{i3}),$$

kde $w_1(v_{ij})$ je marginální hustota pravděpodobnosti j -té kartézské složky vektoru rychlosti i -tého atomu. Opět jednotlivé kartézské složky vektoru rychlosti jsou rovnocenné a tedy i funkce w_1 charakterizující marginální hustotu je pro všechny kartézské složky vektoru rychlosti totožná. Ve výsledku tedy dostáváme, že se systém N částic ideálního plynu skládá z $3N$ nezávislých podsystemů, kde každý odpovídá jedné kartézské složce vektoru rychlosti N částic. Jak víme z předchozího příkladu, entropie takového systému je součtem entropií těchto $3N$ nezávislých podsystemů. Pokud tedy vazbové podmínky představují nezávislé vazby pro jednotlivé kartézské složky vektoru rychlosti N částic, je maximalizace entropie celého podsystemu ekvivalentní maximalizaci entropie každého podsystemu zvlášť. To je ale, jak hned uvidíme, náš případ.

Stačí tedy nalézt rozdělení jedné kartézské složky rychlosti, kterou nyní pro jednoduchost označíme v_j (tedy vynecháváme dále index označující částici). Z faktu, že nádoba s plynem se nehýbe plyne vazbová podmínka $\langle v_j \rangle = 0$. Druhou vazbovou podmínku dostaneme z ekvipartičního teorému, podle kterého má atom plynu při dostatečně vysoké teplotě T střední hodnotu kinetické energie rovnu

$$\langle E_k \rangle = \frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}kT.$$

Protože $v^2 = v_1^2 + v_2^2 + v_3^2$ a všechny kartézské složky jsou rovnocenné, dostaneme pro střední hodnotu kvadrátu jedné kartézské složky rychlosti podmínku

$$\langle v_i^2 \rangle = \frac{kT}{m}.$$

Nejpravděpodobnější rozdělení jedné kartézské složky rychlosti má tedy tvar (viz. Příklad 0.3 pro $\mu = 0$ a $\sigma^2 = \frac{kT}{m}$)

$$w(v_i) = \left(\frac{m}{2\pi kT} \right)^{\frac{1}{2}} \exp \left(-\frac{mv_i^2}{2kT} \right).$$

Rozdělení vektoru rychlosti jednoho atomu je potom

$$w(\vec{v}) = \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} \exp \left(-\frac{mv^2}{2kT} \right),$$

což je nám již dobře známé Maxwelllovo rozdělení vektoru rychlosti molekul ideálního plynu.

Příklad 0.5. Mějme dvě šestistěnné kostky. Opakovaným společným házením kostek jsme zjistili, že stejná čísla padají na kostkách a -krát častěji než různá čísla, kde a je libovolné nezáporné číslo.

- a) Určete nejpravděpodobnější rozdělení výsledků hodu kostkami.
- b) Určete marginální rozdělení výsledků hodu jednotlivých kostek.
- c) Určete kovarianci náhodných veličin X_1 a X_2 , kde X_i je výsledek hodu na i -té kostce.
- d) Dokažte, že kovariance je nulová právě tehdy, když jsou obě kostky nezávislé.
- e) Ukažte, že maximální kovariance (kladná lineární závislost) je dosažena pro $a = 0$.

Návod: Výsledek: a) $p_{ii} = \frac{1}{Z}e^{-\lambda}, p_{ij} = \frac{1}{Z}e^{-a\lambda}$ pro $i \neq j$, kde $e^\lambda = (5a)^{\frac{1}{a-1}}$ a $Z = 6e^{-\lambda} + 30e^{-a\lambda}$, b) hody obou kostek jsou rovnoměrně rozděleny, c) Střední hodnota součinu obou pozorovatelných je $\langle X_1 X_2 \rangle = 7 \frac{a+1.9}{a+1}$. Řešení navazujících problémů je pak již přímočaré.