

Michal Čihák

Ukázka využití počítačů při výuce pravděpodobnosti na gymnáziu (2)

*Učitel matematiky*, Vol. 9 (2001), No. 4, 200–210

Persistent URL: <http://dml.cz/dmlcz/150905>

## Terms of use:

© Jednota českých matematiků a fyziků, 2001

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

## UKÁZKA VYUŽITÍ POČÍTAČŮ PŘI VÝUCE PRAVDĚPODOBNOСТИ NA GYMNAZIU (2)

(Dokončení z minulého čísla)

MICHAL ČIHÁK

### 3. Další pravděpodobnostní pojmy

Předchozí výsledky umožňují řešit celou řadu dalších blízkých úloh. Můžeme se například studentů ptát, jaká je pravděpodobnost  $p_{n,k}$ , že právě  $k$  pánů dostane zpět své klobouky, kde  $0 \leq k \leq n$ .

**Řešení:** Jestliže právě  $k$  pánů má mít svůj klobouk, znamená to:

- Máme  $\binom{n}{k}$  možností pro vybrání těchto  $k$  pánů ze skupiny  $n$  pánů.
- Zbylých  $n - k$  pánů musí mít jiný klobouk než svůj a počet možností, jak toto může nastat, je  $r_{n-k}$ .

Z dvou předchozích tvrzení vyplývá, že počet permutací, které mají  $k$  pevných bodů, je  $\binom{n}{k}r_{n-k}$ , a proto

$$p_{n,k} = \frac{\binom{n}{k}r_{n-k}}{n!}. \quad (13)$$

Získané výsledky velmi dobře poslouží pro ilustraci některých dalších pravděpodobnostních pojmů.

**3.1. Náhodná veličina a její rozdělení.** Nyní si vysvětlíme velmi důležitý pojem tzv. *náhodné veličiny*. Označíme  $X$  počet těch pánů, kteří zachytili svůj klobouk. Veličina  $X$  může „náhodně“ nabývat hodnot  $0, 1, 2, \dots, n$  a proto se nazývá *náhodnou veličinou*. Každé hodnotě je vztahem (13) jednoznačně přiřazena pravděpodobnost, se kterou veličina  $X$  této hodnoty nabývá. Přitom platí

$$\sum_{k=0}^n p_{n,k} = 1,$$

neboť náhodná veličina  $X$  nabude vždy právě jedné z možných hodnot  $0, 1, 2, \dots, n$ .

Takový systém výsledků spolu s jim přiřazenými pravděpodobnostmi nazýváme *pravděpodobnostním rozdělením* náhodné veličiny  $X$ . Pravděpodobnostní rozdělení můžeme vyjádřit například formou tabulky.

Můžeme tedy studentům zadat úlohu vytvořit tabulku všech hodnot náhodné veličiny  $X$  a jim odpovídajících pravděpodobností pro některou konkrétní hodnotu  $n$  (viz např. tabulka 4), nebo ještě lépe několik takových tabulek pro různé hodnoty  $n$  (je vhodné organizovat jako skupinovou práci, každá skupina pracuje na tvorbě jedné tabulky).

$k$	0	1	2
$P(X = k)$	0.366 667	0.375 000	0.166 667

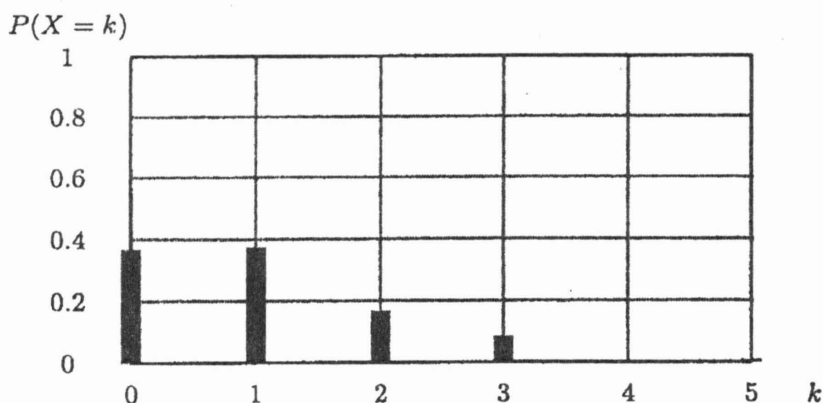
$k$	3	4	5
$P(X = k)$	0.083 333	0	0.008 333

TAB. 4. Pravděpodobnost  $P(X = k)$  pro  $n = 5$ , vypočítaná ze vztahu (13).

Pro větší názornost se pravděpodobnostní rozdělení náhodné veličiny často vyjadřuje grafem. Vyzveme proto studenty, aby se zamysleli, jaký typ grafu zvolit, a pokusili se ho sestrojít. Jednou z možností je například sloupcový graf na obrázku 4.

Zajímavé je, že náhodný jev *právě*  $n - 1$  pánů obdrží svůj klobouk je jev nemožný a jeho pravděpodobnost je tedy nulová (viz tabulka 4 a obrázek 4). Zachytí-li totiž  $n - 1$  pánů svůj klobouk, zbývá již jen jediný klobouk, který musí patřit zbývajícím pánovi a tedy všichni pánové budou mít svůj klobouk.

Nyní vyřešíme tutéž úlohu simulací na počítači. Studentům doporučíme, aby se pokusili modifikovat program 1 tak, aby na výstupu získali tabulku relativních četností výskytu jednotlivých



OBRÁZEK 4. Grafické znázornění rozdělení náhodné veličiny  $X$ .

hodnot  $0, 1, 2, \dots, n$  náhodné veličiny  $X$  při velkém počtu provedených experimentů. Je opět velmi zajímavé porovnat takto získané relativní četnosti s odpovídajícími hodnotami  $p_{n,k}$ . Program 2 je ukázkou, jak je možno uvedenou úlohu řešit.

### Program 2.

```

- - - - - proměnné, konstanty, procedury a funkce - - - - -
n:=10; (* délka permutace *)
poc_exp:=10000; (* celkový počet prováděných experimentů *)
INT i,j;
INT X; (* počet pánů, kteří zachytili svůj klobouk *)
INT vektor[1 TO n]; (* vektor pro náhodné permutace *)
INT abs_cetnost[0 TO n]; (* vektor absolutních četností X *)
REAL rel_cetnost[0 TO n]; (* vektor relativních četností X *)

(* zde budou def. procedur Vloz[] a Zamichej[] - viz. Progr. 1 *)

- - - - - počáteční hodnoty - - - - -
FOR j:= 0 TO n DO
  abs_cetnost[j]:=0; (* explicitní nulování polí *)
  rel_cetnost[j]:=0;
END;
Vloz (vektor[1 TO n]); (* do vektoru vloží čísla 1,2,...,n *)
- - - - - model - - - - -
FOR i:=1 TO poc_exp DO
  Zamichej (vektor[1 TO n]); (* náh. permutace pole [1,2,...,n] *)
  X:=0;

```

```

FOR j:= 1 TO n DO          (* prochází vygenerovanou permutaci *)
  IF vektor[j]=j THEN X:=X+1;
  END; (* najde-li shodu, zvýší čítač pánů se svým kloboukem *)
END;
abs_cetnost[X]:=abs_cetnost[X]+1; (* zvýší čítač veličiny X *)
rel_cetnost[X]:=abs_cetnost[X]/i; (* dělíme počtem experim. *)
VypisPole(rel_cetnost[X]);      (* výpis výsledků do tab. *)
END;

```

K programu dodejme již jen několik drobných poznámek: V systému Famulus není po spuštění programu obsah proměnných a polí inicializován na nulu, ale je jim přiřazena zvláštní nedefinovaná hodnota „?“. Vynulování proměnných je tedy nutno provést explicitně v programu.

Dále není možné použít standardní proceduru DISP pro výpis výsledků, neboť ta nepracuje s poli, ale je nutno pro tyto účely vytvořit vlastní proceduru, která se může jmenovat např. VypisPole().

Z algoritmického hlediska jedinou výraznější změnou oproti programu 1 je již jen použití FOR cyklu místo cyklu WHILE při průchodu polem vektor []. Tentokrát již není možné průchod pole předčasně ukončit a mohou tedy být použity obě varianty, FOR cyklus je však přehlednější.

**3.2. Střední hodnota.** Často se v pravděpodobnostních a statistických úlohách řeší následující problém: *Mnohokrát zopakují experiment, při kterém pánové vyhazují své klobouky. Kolik pánů v průměru zachytí svůj klobouk?* Odpověď na tuto otázku nám může dát výpočet tzv. *střední hodnoty* náhodné veličiny. Její hodnota je definována jako vážený aritmetický průměr všech možných hodnot náhodné veličiny, přičemž každá hodnota je vážena svojí pravděpodobností. Opět ukážeme několik možností, jak střední hodnotu v našem případě určit. To však neznamená, že dále uvedená řešení se dají použít jen pro tuto konkrétní úlohu, naopak, postupy zde uvedené se v různých obměnách používají velmi často. Při řešeních použijeme:

- definici *EX*

- 0–1 náhodné veličiny
- počítačovou simulaci

**3.3. Řešení přímým užitím definice  $EX$ .** Mějme  $n$  pánů a označme  $X$  počet těch pánů, kteří zachytili svůj klobouk. Ten, jak víme, je stejný jako počet pevných bodů v náhodné permutaci  $n$  prvků. Připomeňme, že střední hodnota náhodné veličiny  $X$  je definována takto:

$$EX = \sum_{k=0}^n k P(X = k) = \sum_{k=0}^n k p_{n,k}. \quad (14)$$

Pravděpodobnosti  $p_{n,k} = P(X = k)$  studenti určili již dříve, a proto zbývá jen dosadit do vzorce a pro různé hodnoty  $n$  určit odpovídající střední hodnoty. Studenti budou asi překvapeni, když zjistí, že pro různé hodnoty  $n$  vychází střední hodnota stále stejná a to vždy rovna jedné. Můžeme tedy vyslovit hypotézu, že ať je počet pánů vyhazujících klobouky jakýkoliv, svůj klobouk zachytí v průměru jeden pán. Kdybychom však chtěli uvedenou hypotézu dokázat, asi už s našim postupem výpočtu nevystačíme. Museli bychom totiž pro všechny možné hodnoty  $n$  určit hodnotu  $EX$ . Použijeme tedy obecnější postup.

### 3.4. Řešení užitím 0–1 náhodných veličin.

Nyní uvedeme způsob řešení, který ukazuje, jak výhodné bývá vyjádření náhodné veličiny jako součtu 0 – 1 náhodných veličin. Tento obrat se v pravděpodobnostních úlohách velmi často používá a umožňuje velmi elegantně řešit i značně obtížné úlohy.

Nechť  $X$  je opět počet pánů, kteří zachytili svůj klobouk. Je zřejmé, že  $X$  můžeme vyjádřit jako součet 0–1 náhodných veličin

$$X = U_1 + U_2 + \dots + U_n,$$

kde

$$U_i = \begin{cases} 1 & \text{jestliže } i\text{-tý pán má svůj klobouk,} \\ 0 & \text{jestliže } i\text{-tý pán nemá svůj klobouk.} \end{cases} \quad (15)$$

Dále potřebujeme určit pravděpodobnost jevu, že  $i$ -tý pán má svůj klobouk. Tato pravděpodobnost je rovna, viz (12).

$$P(U_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Nyní vyjádříme  $EU_i$  podle (14)

$$EU_i = 1 \cdot P(U_i = 1) + 0 \cdot P(U_i = 0) = P(U_i = 1) = \frac{1}{n}. \quad (16)$$

Je vidět, proč jsme nepotřebovali určit pravděpodobnost  $P(U_i = 0)$ . Ta se ve vzorci (16) násobí nulou a tedy vypadne. S podobnými zjednodušeními se u 0–1 náhodných veličin setkáváme velmi často. To je také hlavní důvod, proč se takovou měrou v teorii pravděpodobnosti využívají.

Zbývá již jen využít jedné ze základních vlastností střední hodnoty: střední hodnota součtu náhodných veličin je rovna součtu středních hodnot těchto veličin. Proto

$$EX = EU_1 + EU_2 + \dots + EU_n = \frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n} = 1.$$

Naše hypotéza se tedy plně potvrdila.

**3.5. Počítačová simulace.** Opět stačí jednoduchá modifikace našeho programu a můžeme určit střední počet pánů, kteří obdrželi svůj klobouk při mnohonásobném provedení našeho pokusu. Ten můžeme považovat za odhad hledané střední hodnoty náhodné veličiny  $X$ . Zde je jedna z možných modifikací programu:

### Program 3.

```

- - - - - proměnné, konstanty, procedury a funkce - - - - -
n:=10; (* délka permutace *)
poc_exp:=10000; (* celkový počet prováděných experimentů *)
INT i,j;
INT X; (* počet pánů, kteří zachytili svůj klobouk *)
INT vektor[1 TO n]; (* vektor pro náhodné permutace *)
INT suma; (* kumulativní součet hodnot veličiny X *)
REAL stred_hodnota; (* střední hodnota veličiny X *)

```

```

(* zde budou def. procedur Vloz[] a Zamichej[] - viz. progr. 1 *)

- - - - - počáteční hodnoty - - - - -
suma:=0;
Vloz (vektor[1 TO n]);      (* do vektoru vloží čísla 1,2,..,n *)
- - - - - model - - - - -
FOR i:= 1 TO poc_exp DO
  Zamichej (vektor[1 TO n]); (* náh. permutace pole [1,2,..,n] *)
  X:=0;
  FOR j:=1 TO n DO          (* prochází vygener. permutaci *)
    IF vektor[j]=j THEN X:=X+1;
    END; (* najde-li shodu, zvýší čítač páňů se svým kloboukem *)
  END;
  suma:=suma+X;
  stred_hodnota:=suma/i;    (* dělí počtem experimentu *)
  DISP;                    (* zobrazí tab. a graf relat. četnosti *)
END;

```

Algoritmus je obdobný jako u programu 2, jediný rozdíl spočívá v použití proměnné *suma*, do které jsou postupně přičítány počty pevných bodů zjištěné v náhodně vygenerovaných permutacích.

#### 4. Rozptyl

Představme si, že bychom řešili následující problém. Víme, že střední počet páňů, kteří zachytí svůj klobouk, je roven jedné bez ohledu na počet páňů. Jaké výsledky však můžeme obvykle očekávat? Budou výsledky spíše „rozptýlenější“, tzn. budou se mezi nimi často objevovat i hodnoty od střední hodnoty výrazně vzdálené, nebo budou výsledky ke střední hodnotě těsně přimknuté a vzdálenější hodnoty se budou vyskytovat jen velmi vzácně? Odpověď na tuto otázku může dát výpočet tzv. *rozptylu* náhodné veličiny.

Výpočet rozptylu je již problémem, který vyžaduje znalost některých dalších pojmů z teorie pravděpodobnosti a výrazněji tak překračuje rámec středoškolských znalostí. Studenti musí znát nejen pojem rozptylu a jeho základní vlastnosti, navíc musí znát i pojem tzv. *kovariance* náhodných veličin.



Rozptyl náhodné veličiny  $X$  je definován takto

$$\text{var}X = E(X - EX)^2, \quad (17)$$

po úpravě a užitím vlastností střední hodnoty dostaneme

$$\begin{aligned} \text{var}X &= E(X^2 - 2XEX + (EX)^2) \\ &= EX^2 - 2(EX)^2 + (EX)^2 = EX^2 - (EX)^2. \end{aligned} \quad (18)$$

Kovariance libovolných náhodných veličin  $X, Y$  je definována

$$\text{cov}(X, Y) = E(X - EX)(Y - EY), \quad (19)$$

a po obdobné úpravě

$$\text{cov}(X, Y) = EXY - EXEY. \quad (20)$$

V úloze chceme určit  $\text{var}X$ , kde  $X$  je počet těch pánů, kteří zachytili svůj klobouk. Veličinu  $X$  opět vyjádříme jako součet 0-1 náhodných veličin  $U_1, \dots, U_n$  definovaných v (15). Pro rozptyl součtu  $n$  náhodných veličin můžeme s využitím (17) a (19) psát

$$\begin{aligned} \text{var}X &= \text{var}(U_1 + U_2 + U_3 + \dots + U_n) \\ &= E[(U_1 + \dots + U_n) - E(U_1 + \dots + U_n)]^2 \\ &= E[(U_1 - EU_1) + (U_2 - EU_2) + \dots + (U_n - EU_n)]^2 \\ &= \text{cov}(U_1, U_1) + \text{cov}(U_1, U_2) + \dots + \text{cov}(U_1, U_n) \\ &\quad \vdots \\ &\quad + \text{cov}(U_n, U_1) + \text{cov}(U_n, U_2) + \dots + \text{cov}(U_n, U_n) \\ &= \sum_{i=1}^n \text{var}U_i + \sum_{1 \leq i \neq j \leq n} \text{cov}(U_i, U_j), \end{aligned} \quad (21)$$

neboť

$$\text{cov}(U_i, U_i) = E(U_i - EU_i)(U_i - EU_i) = E(U_i - EU_i)^2 = \text{var}U_i$$

pro každé  $i$ .

Nyní opět využijeme velmi výhodné vlastnosti 0-1 náhodných veličin. Platí, že  $U_i^2 = U_i$  a proto můžeme pro rozptyl  $U_i$  podle (18) psát

$$\text{var}U_i = EU_i^2 - (EU_i)^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2}.$$

Pro kovariance náhodných veličin  $U_i, U_j$ , kde  $i \neq j$ , je podle (20)

$$\text{cov}(U_i, U_j) = EU_iU_j - EU_iEU_j.$$

Nyní potřebujeme určit  $EU_iU_j$ . To se nám podaří velmi jednoduše díky výhodným vlastnostem 0-1 náhodných veličin. Pro názornost je vhodné vypsát všechny možné kombinace hodnot veličin  $U_i$  a  $U_j$  a jim odpovídající hodnoty náhodné veličiny  $U_iU_j$  v tabulce 5.

	$U_j$	
$U_i$	0	1
0	0	0
1	0	1

TAB. 5. Všechny možné hodnoty, kterých může nabývat náhodná veličina  $U_iU_j$ .

Zbývá přiřadit každé kombinaci hodnot z tabulky odpovídající pravděpodobnost a můžeme psát

$$\begin{aligned} EU_iU_j &= 0 \cdot 0 \cdot P(U_i = 0, U_j = 0) + 0 \cdot 1 \cdot P(U_i = 0, U_j = 1) \\ &\quad + 1 \cdot 0 \cdot P(U_i = 1, U_j = 0) + 1 \cdot 1 \cdot P(U_i = 1, U_j = 1) \\ &= P(U_i = 1, U_j = 1). \end{aligned} \tag{22}$$

Je zřejmé, že

$$P(U_i = 1, U_j = 1) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$

a proto z (22) obdržíme

$$EU_i U_j = \frac{1}{n(n-1)}.$$

Nyní již vyjádříme kovarianci

$$\text{cov}(U_i, U_j) = EU_i U_j - EU_i EU_j = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}.$$

Vzpomeňme si nyní na vzorec (21). Výraz  $\text{cov}(U_i, U_j)$  je zde obsažen  $n(n-1)$  krát. Celkem tedy dostáváme

$$\begin{aligned} \text{var} X &= \text{var} \sum_{i=1}^n U_i = \sum_{i=1}^n \text{var} X_i + \sum_{1 \leq i \neq j \leq n} \text{cov}(U_i, U_j) \\ &= n \frac{n-1}{n^2} + n(n-1) \frac{1}{n^2(n-1)} = 1. \end{aligned}$$

Rozptyl je tedy stejně jako střední hodnota roven jedné bez ohledu na počet pánů.

#### LITERATURA:

- [Blom] Blom, G., Holst, L., Sandell, D., *Problems and Snapshots from the World of Probability*, Springer-Verlag, New York, 1994.  
 [Isaac] Isaac R., *The Pleasures of Probability*, Springer-Verlag, New York, 1995.  
 [Most] Мостеллер, Ф., *Пятьдесят занимательных вероятностных задач с решениями*, Наука, Москва, 1971.

---

Autor děkuje doc. RNDr. Jaromíru Antochovi, CSc. z KPMS MFF UK za cenné rady, náměty a připomínky, které mu po celou dobu přípravy článku ochotně poskytoval.

- [Ross] Ross, Sheldon M., *A First Course in Probability*, Macmillan Publishing, New York, 1988.
- [Sim] Simon, J. L., Atkinson, D., Shevokas, C., *Probability and statistics: Experimental results of radically different teaching method*, The American Mathematical Monthly, Vol. 83 (1976), No. 9, 733–739.
- [Step] Štěpánová, I. a Štěpán, J., *Osm úloh o kombinatorické pravděpodobnosti*, Matematika – fyzika – informatika 3 (1993–1994), 113–119.
- [Hejn] Hejný, M. a kol., *Teória vyučovania matematiky 2*, SPN, Bratislava, 1990.
- [Fam] Dvořák, L., Ledvinka, T., Sobotka, M., *Famulus 3.1 – příručka uživatele*, Praha, 1991.

*Mgr. Michal Čihák*

*doktorand Katedry pravděpodobnosti a matematické statistiky MFF UK*

*Sokolovská 83, 186 75 Praha 8*

*email: mcihak@cmail.cz*