

Proč se vlastně učit a studovat statistiku

Jeden z možných důvodů je záměrná nebo i nezáměrná manipulace s daty.
Příklad, povodně.

Zdrojová data:

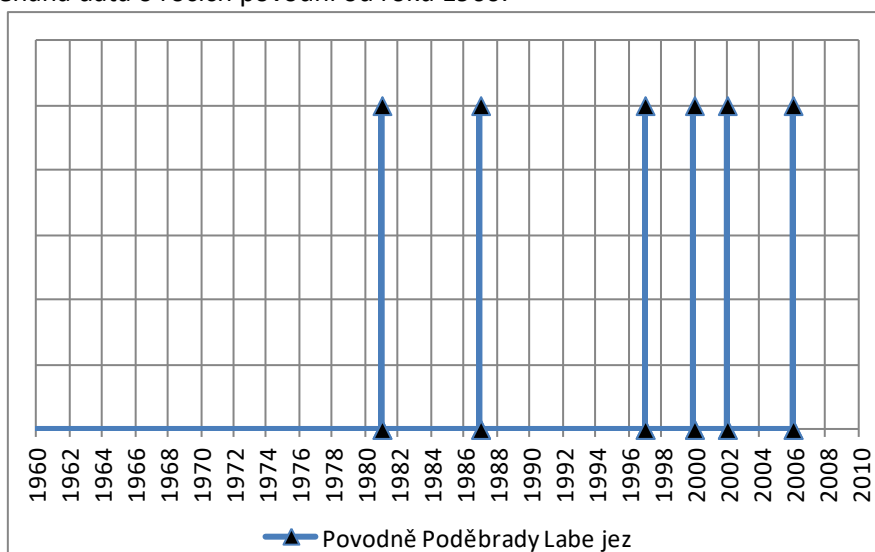


Poděbrady jez.

Pozorované roky jednotlivých povodní:

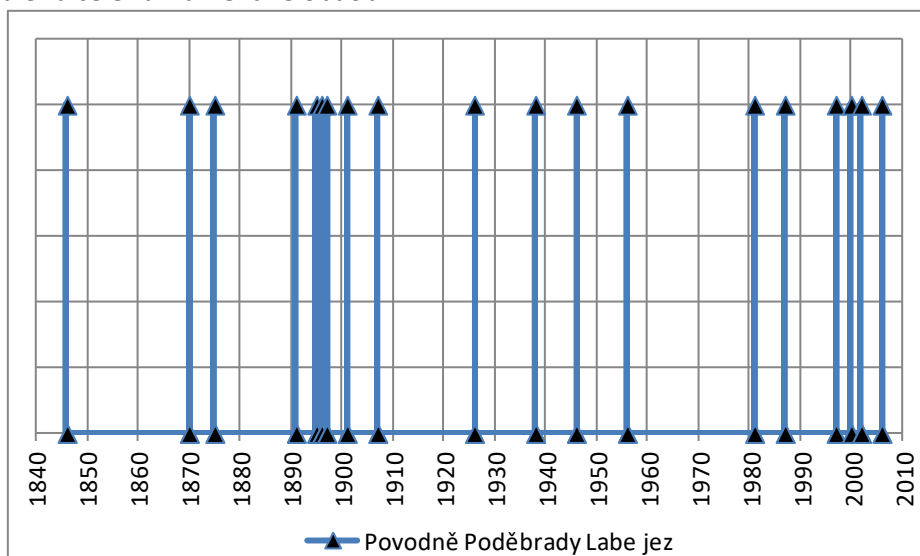
Minimum	1
Průměr	8,89
Medián	6
Maximum	25
StD	7,53
Počet	18
Rok	Doba mezi
1846	
1870	24
1875	5
1891	16
1895	4
1896	1
1897	1
1901	4
1907	6
1926	19
1938	12
1946	8
1956	10
1981	25
1987	6
1997	10
2000	3
2002	2
2006	4

Jinak zaznamenaná data o rocích povodní od roku 1960.

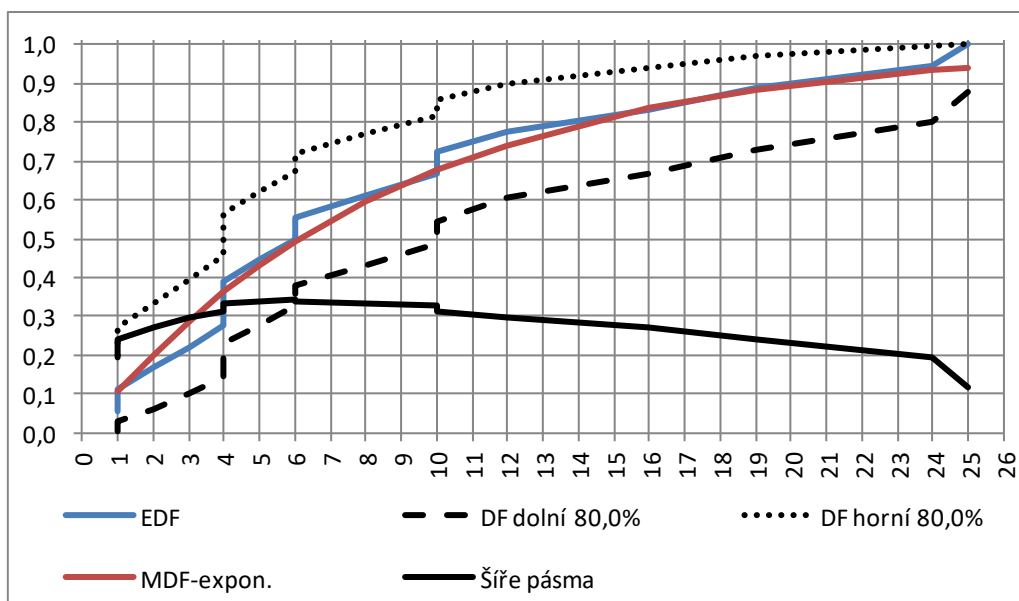


Pravdivý, ale zavádějící a neúplný závěr z tohoto obrázku: **V poslední době jsou povodně daleko častější než dříve.**

Tatáž data, ale za celé zaznamenané období:

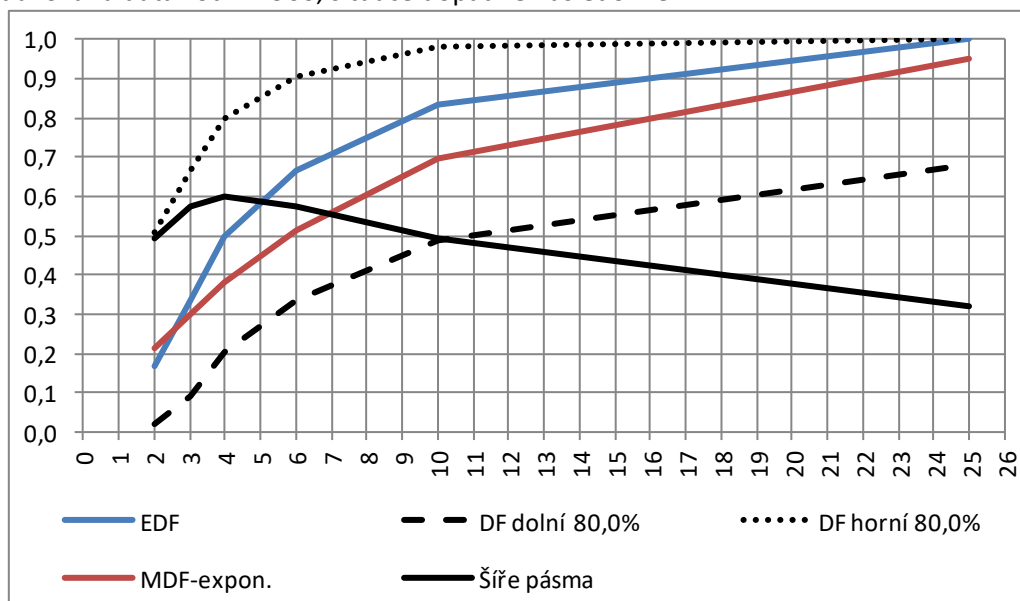


Z tohoto obrázku již nelze, tak snadno, učinit předchozí závěr! Spíše se budeme klonit k závěru: **Jedná se o stacionární proces s exponenciálním rozdělením dob mezi povodněmi.** To také naznačuje empirická distribuční funkce s 80% pásem spolehlivosti pro skutečnou distribuční funkci, včetně modelové distribuční funkce exponenciálního rozdělení.

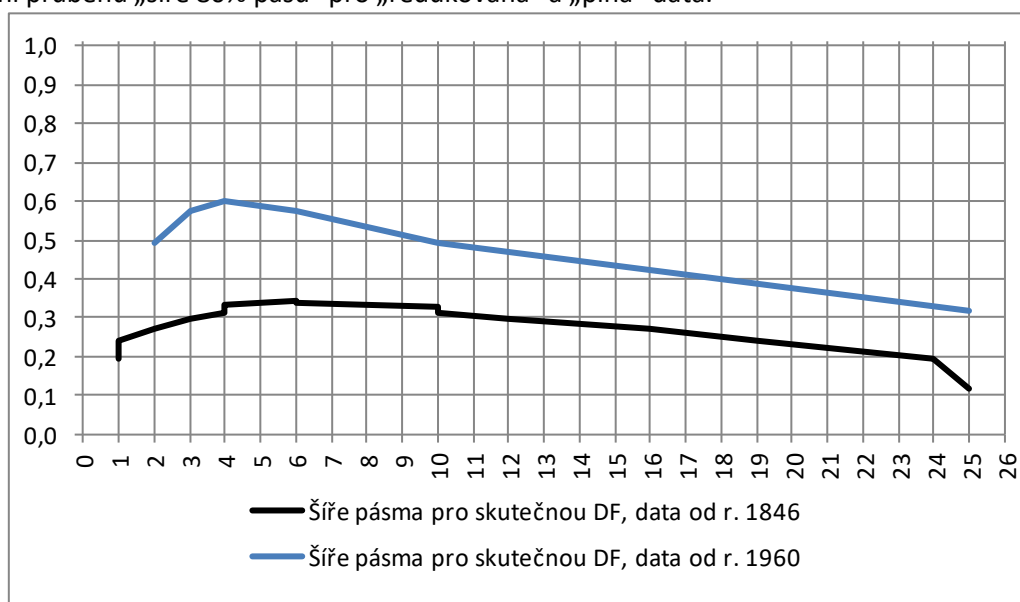


Důležitým ukazatelem je zde šíře pásma pro danou „míru spolehlivosti“ (zde 80%), ta je ukazatelem pochybností o (zde příliš viditelné) shodě mezi modelem a pozorovanými daty (zde prezentovanými empirickou distribuční funkcí).

Pro „redukována data“ od r. 1960, situace dopadne následovně:



Srovnání průběhů „šíře 80% pásu“ pro „redukována“ a „plná“ data.



Poznámka a problém: Data jsou v celých rocích, správněji by měla být vyjádřena i v částech roků.

Kriteriem pro to, co si máme myslet o jednotlivých datových souborech a z nich odhadnutých modelech je zde šíře pásma pro skutečnou DF.

V demonstrované úloze je vše jasné na první pohled. Leč s jinými daty (i částečně zamlčenými) může být statistický aparát „dobrým měřidlem“ toho jaké závěry o výběru dat a jejich zpracování máme přijímat. Z druhé strany výběrem dat a s nimi spojeného modelu lze vyvolávat „předem daný a žádoucí (i když třeba nesprávný) dojem“. Známé z mediálních tvrzení typu „Vědci z univerzity xxxxx dokázali, že yyyy“.

Další problém: nedostatečné znalosti pravděpodobnostních a statistických charakteristik některých datových souborů.

Téměř trvalá a periodická story v českých médiích: naposled po 6.9.2013 (<https://www.czso.cz/csu/czso/prumerne-mzdy>). ČSÚ nějaký čas (cca 2 měsíce) po každém kalendářním čtvrtletí zveřejňuje statistiku, v daném čtvrtletí, dosahovaných mezd. Pokaždé se v médiích objeví zpráva (bez dalšího vysvětlení), že 2/3 pracovníků mají **podprůměrnou** mzdu a naznačuje se, že se jedná o cosi nesprávného. Někteří takzvaní „odborníci doporučují“, že by s tím mělo být něco děláno (vláda?).

Příklad:

komentare.iHned.cz

úterý, 12. října 2004

NÁZORY

HOSPODÁŘSKÉ NOVINY

Jsou další platy nadbytečné?

Dohadovací řízení kolem platů vyústilo v nečekaně košatý problém. Zdá se, že ho v dohledné době nezvládneme. Nejen u zaměstnanců. Podnikovou mikroúroveň počínaje a státní makroúrovň konče bude třeba leccos ještě řešit.

Dosud přijaté závěry místy připomínají účet uzavřený bez hostinského. Výklad «dole» je i ten, že ve svých rozhodnutích zasahuje vláda vedoucím do jejich kompetencí. Ti mají teď v mnohém i více přivřeně dveře. Pomíneme-li výši platů (někde stále až skandálně nízkou), potom si omezení plynoucí z jejich přerozdělení a rozpouštění asi vyberou svoji daň.

Efektivní motivační přístupy stále nejsou naší parketou. Vedoucí mají i nadále problémy, a to nejen s finanční motivací. Loňské ankety ke 13. a 14. platům ukázaly na podivuhodnou jednotu v českých podnicích. Averse vůči více výplatám než má rok měsíců byla snad stoprocentní. Expresivní výrazy typu «nejsem padlí na hlavu» nebyly výjimkou.

Vypadá to, že vedoucí jsou zjedno v názoru, že jediným správným motivačním motorem v podnicích je, když plat kopíruje výkon. To by samo o sobě nebylo vůbec špatné (a bylo by úplně nejlepší, být tomu tak opravdu všude), kdyby nešlo vesměs o jediné oce-

nění, kterého se pracovníkům za dobrou práci dostane. Problémem je, že vždy se najde nějaká firma, která může nabídnout víc...

Není proto divu, že uspěje každý dobře postavený inzerát nebo «laň», který se namane (o «lovcech mozků» ani nemluvě), protože lidi kromě peněz v takovém podniku nic jiného nedrží («...do nestabilních pracovníků se přece nevyplatí nic investovat...»).

Vedoucí či majitel zřejmě netuší, že právě takto zjednodušeně chápanou motivační politikou připravuje firmu o její nejlepší pracovníky. Jen pomalu přibývá těch, kteří růstu výkonnosti svých dobrých pracovníků pomáhají dalším vzděláváním a postupem počínaje a různými výhodami či věcnými bonusy konče.

Riziko odchodů

Problém fluktuace by možná tolik nebolel, kdyby se už nezačal týkat i úzkoprofilových profesí, na nichž je jednoznačně závislá prosperita mnohých podniků. Odchod ke konkurenci, popřípadě do jiné profese, představuje leckdy nevratnou ztrátu.

A bude hůř. Stále nekončící «zeštíhlovací» tendence směřují k tomu, že v podnicích budou postup-

ně ponecháváni jen ti skutečně nenahraditelní («znalostní» pracovníci) a ostatní se stanou v různé formě externími (zkrácený úvazek, pronájem pracovníků, outsourcing).

Vyplatilo by se udělat něco pro loajalitu. Loajální pracovník je cenný pracovník. Ten k firmě patří a ona k němu (protože si ho zasloužila díky přístupu svých ve-

doucích i pracovními a sociálními podmínkami). Svoji firmu na veřejnosti nepomlouvá, neplýtvá, nekrade apod. a nedovolí to ani jiným (šetří náklady). A velmi, velmi zvažuje, stojí-li nabízené (za tím většinou nepřelí-
liš) větší peníze jinde za riziko možné budoucí nepohody.

Zásluha podniků, které si «produkuji» věrné zaměstnance je i v tom, že je dobře znají a respektují jejich potřeby. A tak třeba ví, že «zabodují», když například pomohou životní rodiny zvednout sebevědomí a přidat mu doma na důležitosti tím, že před dovolenou může citelně přispět do rodinného rozpočtu. Protože statistici opět potvrdili, že většina pracovníků u nás je stále svým platem pod celorepublikovým průměrem, nejde o okrajovou záležitost ani náhodou.

Zaměstnanců (státních úředníků, soudců, lékařů aj.), kteří zpravidla berou násobek (až násobky) průměrných platů, se takové úvahy na první pohled netýkají. Ale jen na první pohled. Konečný důsledek je však stejný jako v předchozím případě, protože navíc obdržené peníze se (ale zase – jen spolu s dobrými přístupy vedení) projeví ve spokojenosti, a tedy i v loajalitě.

Vážné dopady

Rozhodnutí vlády zatím nejsou příliš šťastná. Spolu s omezenými znalostmi vedoucích, týkající se motivačních technik, se dopady takového souběhu ukazují jako vážné a nesou sebou dost daleko sahající ekonomické komplikace. Třinácté a čtrnácté platy jsou příležitostí k posunu, k účinnějším motivačním technikám i stabilizaci dobrých pracovníků.

I když se jedná o záležitost více než složitou a obsáhlou, v níž asi nikdy nebudeme vědět a umět všechno, neměli bychom platy navíc smahem zavrhnout; dobří pracovníci si je zaslouží a měli by je dostávat, i vše, co k tomu patří. Naopak bychom měli v podnicích a dalších organizacích alespoň ty největší mezery tohoto druhu urychleně zaplnit a «nahofe» učiněná rozhodnutí by tomu propůjčit překážet už neměla.

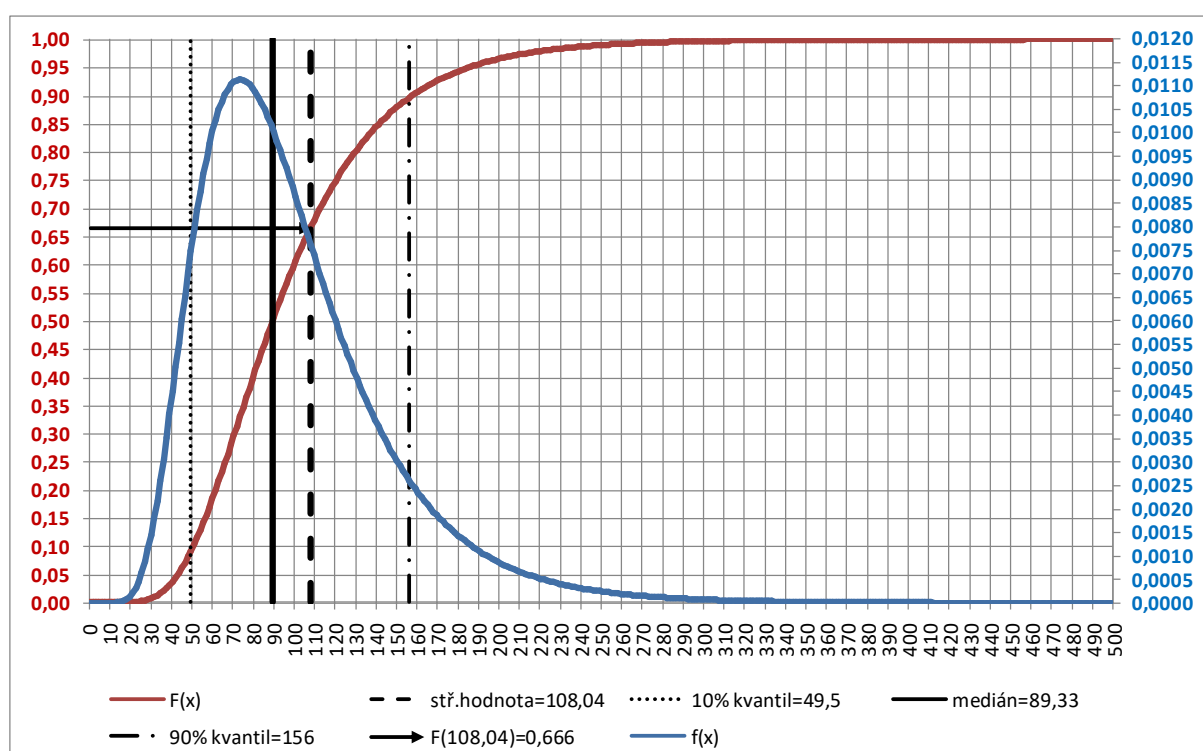
Autor předkládá na Západočeské univerzitě v Plzni a Chebu

Problém je v tom, že zveřejněné číslo je běžnou populací chápáno jako jakýsi střed. Tedy půl pod a půl nad. Běžná populace vlastně nechce průměr jako odhad střední hodnoty ale výběrový medián jako odhad mediánu rozdělení (volněji a nepříliš přesně: medián je číslo, které množinu možných hodnot dělí na dvě stejně pravděpodobné části). ČSÚ již poměrně dlouhou dobu zveřejňuje (výběrový) medián, včetně vysvětlení o co jde. Ale oné běžné populaci je slůvko medián neznámé, tedy nepřátelské. U symetrických rozdělení se medián a střední hodnota shodují (pokud existují), u nesymetrických ne.

Podstata problému je v tom, že rozdělení platů a mezd se přibližně řídí log-normálním rozdělením s hustotou:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\lg(x)-\mu)^2}{2\sigma^2}}$$

s parametry $median = e^\mu$, $E\{x\} = e^{\mu + \frac{\sigma^2}{2}} = median * e^{\frac{\sigma^2}{2}}$ a se šikmostí $(e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$!!!!!



Rozdělení hodinových výdělků.

Ve 2. čtvrtletí 2013 vzrostla průměrná hrubá měsíční nominální mzda na přepočtené počty zaměstnanců v národním hospodářství proti stejnému období předchozího roku o 1,2 %, reálně poklesla o 0,3 %. Medián mezd činil 20 944 Kč.

Ve 2. čtvrtletí 2013 činila **průměrná hrubá měsíční** nominální mzda^{*)} (dále jen „průměrná mzda“) na přepočtené počty zaměstnanců v národním hospodářství celkem **24 953 Kč**, což je o 297 Kč (1,2 %) více než ve stejném období roku 2012. Spotřebitelské ceny se zvýšily za uvedené období o 1,5 %, reálně se tak mzda snížila o 0,3 %. Objem mezd zůstal na stejné úrovni jako ve 2. čtvrtletí minulého roku, počet zaměstnanců poklesl o 1,2 %.

V podnikatelské sféře se průměrná mzda zvýšila nominálně o 1,1 %, reálně se snížila o 0,4 %, v nepodnikatelské sféře vzrostla nominálně o 1,9 %, reálně o 0,4 %.

Proti předchozímu čtvrtletí se průměrná mzda ve 2. čtvrtletí 2013 po očištění od sezónních vlivů zvýšila o 0,4 %.

Medián mezd (20 944 Kč) vzrostl proti stejnému období předchozího roku o 1,3 %, u mužů činil 22 769 Kč, u žen byl 18 836 Kč.

Osmdesát procent zaměstnanců pobíralo mzdu mezi 9 986 Kč a 40 301 Kč.

V 1. pololetí 2013 dosáhla průměrná mzda výše 24 503 Kč, v meziročním srovnání činil přírůstek 101 Kč (0,4 %). Spotřebitelské ceny se zvýšily za uvedené období o 1,7 %, reálně se mzda snížila o 1,3 %.

^{*)} Průměrná hrubá měsíční mzda je podíl mzdových prostředků (vč. příplatků za přesčas, odměn, náhrad mzdy atd.) připadající na jednoho zaměstnance za měsíc. Nevypovídá tedy o výplatě jednoho konkrétního zaměstnance. Výsledky strukturálních výdělkových statistik, kdy jsou k dispozici údaje o výdělích jednotlivých zaměstnanců, uvádějí, že zhruba dvě třetiny zaměstnanců mají mzdu nižší než je celostátní průměr.

Medián představuje hodnotu mzdy zaměstnance uprostřed mzdového rozdělení; to znamená, že polovina hodnot mezd je nižší a druhá polovina je vyšší než medián. Na rozdíl od průměrné mzdy, která je vypočtena na základě podkladů z podnikového výkaznictví, je medián nutně odvozovat ze statisticko-matematického modelu na podkladě výběrového šetření, protože podnikové výkazy obsahují jen agregované údaje za celý podnik nebo organizaci.

Nelze zapomenout, že z hrubé mzdy jsou zaměstnavatelem za zaměstnance ještě odvedeny příslušné částky na zdravotní pojištění, sociální zabezpečení a zálohy na daně z příjmu, zaměstnanci je vyplácena mzda čistá. <http://www.czso.cz/csu/csu.nsf/informace/cpmz090613.doc>