

Protokol SME

Daniel Wohlrath

5. srpna 2020

Data 13

Čtyřiačtyřicet dětí – nový rekord – se narodilo během dvaceti čtyř hodin 18. prosince, 1997 v nemocnici „Mater Mothers’ Hospital“ v Brisbane v Austrálii. Pro každé dítě byly zaznamenány údaje o čase narození, pohlaví a porodní váze. Data byla zveřejněna v Brisbaneském nedělníku „The Sunday Mail“ 21. prosince, 1997.

1 Deskriptivní statistika

Tabulka 1 obsahuje hodnoty základních statistik pro porodní váhu dětí v závislosti na pohlaví a čase narození. Z uvedených hodnot průměrů všech dětí, kluků a holek narozených během celého dne (tj. první až třetí řádek) můžeme vypočítat, že kluci mají v průměru vyšší porodní hmotnost, než holky. Navíc při porovnání příslušných průměrů s mediány a směrodatnými odchylkami je zřetelné, že porodní váha u holek má větší rozptyl, případně došlo při měření k výskytu více odlehlých hodnot, které ovlivnily tyto data.

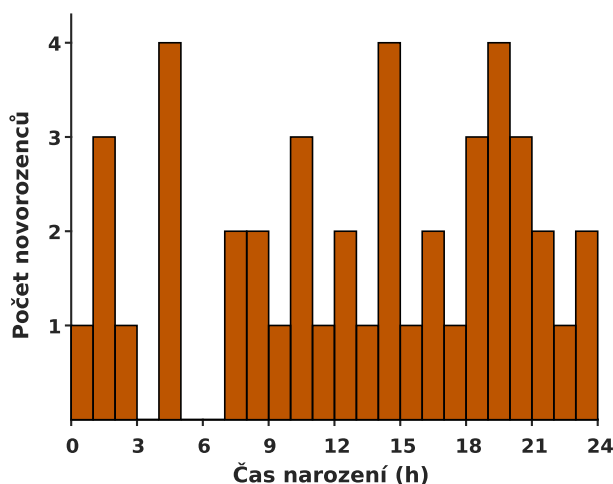
	průměr (g)	směrodatná odchylka (g)	medián (g)	min (g)	max (g)	počet v kategorii
všechny děti	3275.95	528.03	3404	1745	4162	44
kluci	3375.31	428.05	3404	2121	4162	26
holky	3132.44	631.58	3381	1745	3866	18
kluci (0-12h)	3364.60	316.93	3450	2846	3838	10
kluci (12-24h)	3382.00	494.96	3404	2121	4162	16
holky (0-12h)	3022.13	761.91	3271	1745	3837	8
holky (12-24h)	3220.70	531.50	3429	2184	3866	10
všechny děti (0-12h)	3212.39	568.22	3357	1745	3838	18
všechny děti (12-24h)	3319.96	505.06	3417	2121	4162	26

Tabulka 1: Tabulka znázorňující základní statistické údaje pro naměřená data, která jsou rozdělena do kategorií podle pohlaví a času narození.

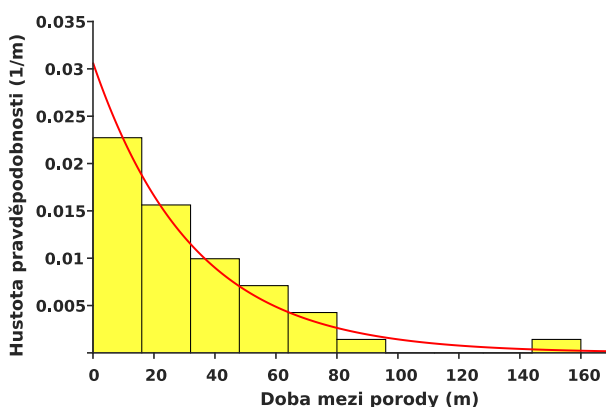
Obrázky 1 a 2 prezentují pomocí histogramů počet porodů v jednotlivých hodinách celého dne a časové intervaly mezi jednotlivými porody. Střední hodnota doby mezi dvěma porody je 32

minut a 36 sekund. Z obrázku 2 je navíc zřejmé, že data velmi dobře odpovídají exponenciálnímu rozdělení, tento fakt bude ověřen v další části tohoto textu.

Uskutečnění porodu může, ale nemusí být nezávislé na ostatních porodech. Jelikož zveřejněná data byla dle zdroje rekordní, je možné, že porodní sestry a porodníci byly v tento den velmi vytíženy. Nemohlo tedy během krátké doby dojít k více porodům, než je k dispozici lékařů, a proto mohly být porody ovlivňovány různými léky, například za účelem jejich oddálení do doby, než se lékaři dostaví z jiného porodu. Jelikož takové údaje k dispozici nemáme, předpokládejme v další analýze, že k takovému ovlivňování porodů nedocházelo, a tedy jednotlivé časy porodů jsou nezávislé události.



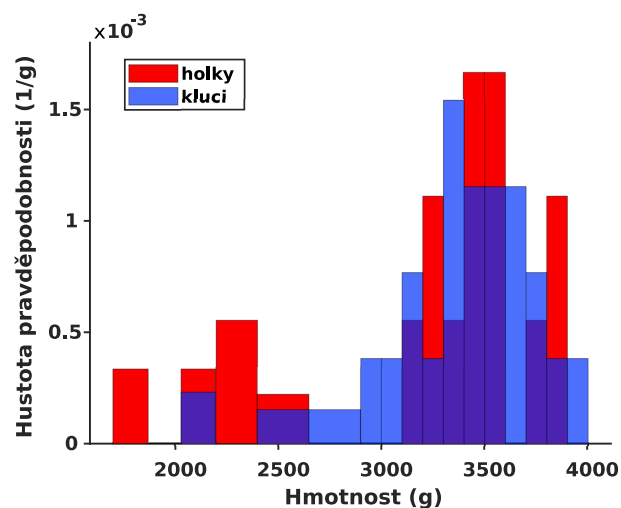
Obrázek 1: Četnost novorozenců narozených v jednotlivých hodinách



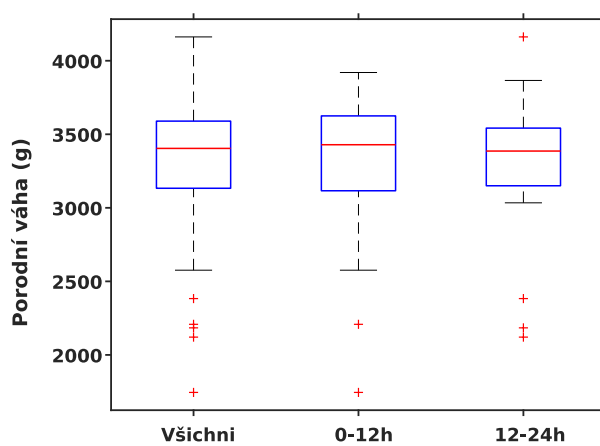
Obrázek 2: Histogramový odhad hustoty pravděpodobnosti (žluté sloupce) a hustota pravděpodobnosti exponenciálního rozdělení se střední hodnotou rovnou MLE (červená plná čára) pro dobu čekání před dalším porodem.

Z obrázku 3 je zřejmé, že holky se rodí při nižších porodních vahách častěji, než kluci. Tento fakt lze matematicky interpretovat způsobem, že levý chvost hustoty pravděpodobnosti porodní hmotnosti kluků klesá rychleji, než levý chvost hustoty pro holky.

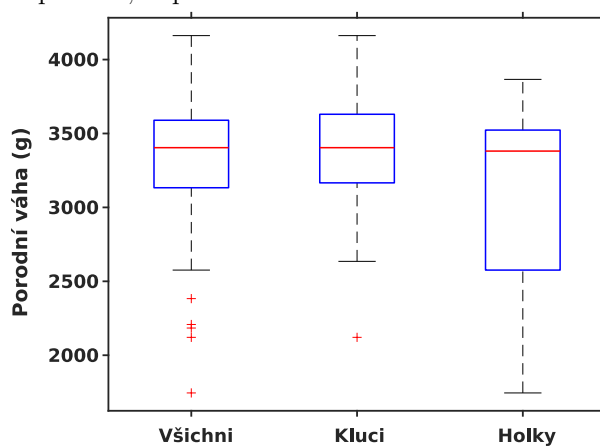
Boxplot na obrázku 4a odpovídá úvaze, že porodní váha novorozenců pravděpodobně nebude záviset na čase porodu. Na druhou stranu boxplot na obrázku 4b poukazuje na fakt, že průměrná porodní hmotnost holek je nižší, než průměrná porodní hmotnost kluků, přičemž mediány jednotlivých kategorií jsou přibližně stejné.



Obrázek 3: Histogramový odhad hustoty pravděpodobnosti pro váhu novorozených kluků (modře) a holek (červeně).



(a) Boxplot pro porodní hmotnost dětí narozených dopoledne, odpoledne a během celého dne.



(b) Boxplot pro porodní hmotnost kluků, holek a společně obou pohlaví.

Obrázek 4: Boxploty pro porodní hmotnost dětí v různých kategoriích

2 Vlastní analýza

Pro všechny následující testy byla zvolena hladina významnosti testu $\alpha = 0,05$. V uvedených Pearsonových χ^2 testech dobré shody jsme slučovali sousední buňky, pokud počet pozorování v jednotlivých buňkách nebyl větší než 5. Pokud bylo použité jiné pravidlo slučování, je na tuto skutečnost upozorněno. Provedené výpočty byly uskutečněny v programovém prostředí MATLAB.

2.1 Testy dobré shody

Nejdříve testujeme hypotézu, zda počet narozených holek, resp. kluků do prvního narození kluka, resp. holky má geometrické rozdělení. Tedy v prvním jmenovaném případě testujeme hypotézu

$$H_0 : F = \text{Geom}(p) \quad \times \quad H_1 : \text{jinak},$$

kde F je distribuční funkce příslušná uvedené náhodné veličiny, p je pravděpodobnost, že se narodí kluk (úspěch) a $1 - p$ je pravděpodobnost narození holky (neúspěch). Definiční obor této veličiny jsme rozdělili na 4 boxy se středy v bodech 0, 1, 2, 3 a při použití Pearsonova χ^2 testu jsme v tomto případě obdrželi hodnotu pearsonovi statistiky $2,793 < \chi^2_{0,95}(2)$ a p-hodnotu 0,248. Proto hypotézu nezamítáme. Podobně jsme otestovali pomocí Pearsonovo χ^2 testu hypotézu druhou variantu. Při stejném nastavení hladiny významnosti testu a rozdělení definičního oboru na 5 boxů se středy v bodech 0, 1, 2, 3, 4 jsme obdrželi p-hodnotu 0,08 a proto ani nyní hypotézu nezamítáme. Grafické znázornění příslušných distribučních funkcí je zobrazeno na obrázcích 5a a 5b.

Dále testujeme hypotézu, zda doba mezi porody má exponenciální rozdělení. Grafická vizualizace tohoto testu je znázorněna na obrázcích 2 a 6. Provedli jsme nejdříve Pearsonův χ^2 test, který hypotézu nezamítnul s p-hodnotou 0,851, která svědčí o velmi dobré shodě dat s teoretickým modelem. Dále jsme použili Lillieforsův test, jež také nezamítnul nulovou hypotézu, a to s p-hodnotou 0,247.

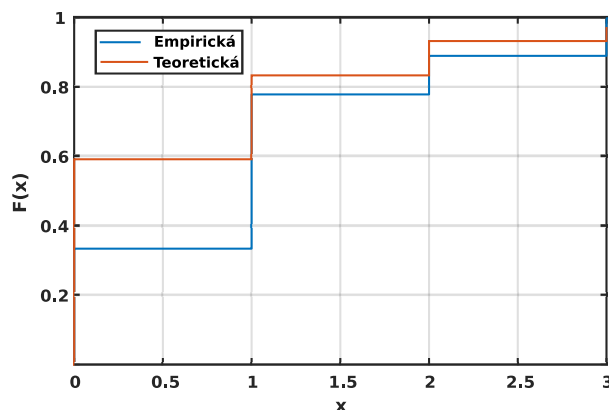
Dále uvažujeme hypotézu, že počet novorozenců za hodinu pro každou hodinu má Poissonovo rozdělení. Odhadnutá hodnota parametru je $\lambda = 1,833$ a shodu dat s modelem jsme ověřili grafickým porovnáním distribučních funkcí, viz obrázek 7 a Pearsonovým χ^2 testem, který hypotézu nezamítnul s p-hodnotou 0,958.

2.2 Testování hypotéz

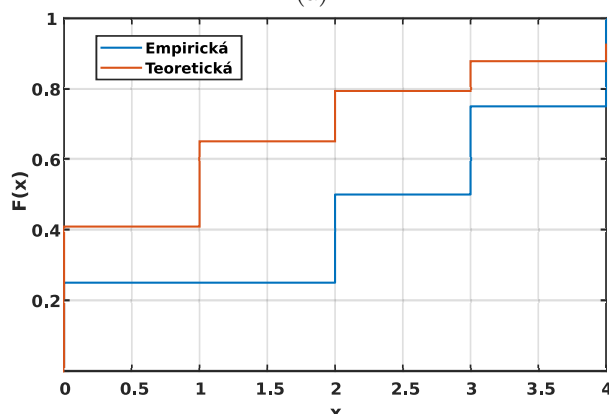
Hlavní otázkou je, zda porodní váha závisí na pohlaví. Nejdříve otestujeme, zda rozdělení odpovídající porodním vahám pro kluky a pro holky je Gaussovo.

Použitím Pearsonova testu na hypotézu, zda porodní váha kluků má normální rozdělení s odhadnutými parametry $\mu = 3375,31$ a $\sigma = 428,05$ jsme obdrželi p-hodnotu 0,42, přičemž jsme upravili pravidlo slučování boxů při výskytu malého počtu pozorování, a to na dvě. Hypotézu jsme nezamítli ani použitím Lillieforsova testu, přičemž jeho p-hodnota byla 0,103.

Analogicky jsme postupovaly při testování hypotézy pro holky, kde $\mu = 3132,44$ a $\sigma = 631,58$. Použili jsme Pearsonův test, opět s upraveným slučovacím pravidlem (dvě pozorování) a výslednou p-hodnotou 0,135, tedy tento test hypotézu nezamítnul. Avšak Lillieforsův test hypotézu normality zamítnul, a to s p-hodnotou 0,028. Poté byl proveden Lillieforsův test znovu, ale za nulovou hypotézu jsme uvažovali, že rozdělení porodní váhy holek pochází z Weibullovo rozdělení s parametry $A = 3371,18$ a $B = 6,86$. Hypotéza zamítnuta nebyla, jelikož jsme obdrželi p-hodnotu 0,137. Podobně pak ani Pearsonův test hypotézu Weibullovo rozdělení nezamítnul s p-hodnotou 0,274. Srovnání odhadnutých hustot pravděpodobnosti normálního rozdělení je na obrázku 8.



(a)



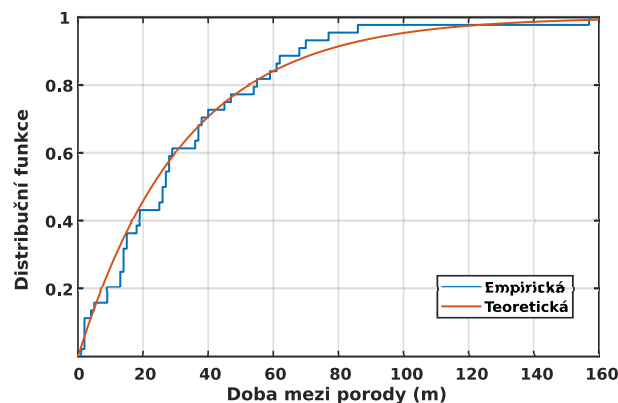
(b)

Obrázek 5: Porovnání empirické a teoretické distribuční funkce při testování hypotézy, zda počet narozených holek do prvního narození kluka (a) následuje geometrické rozdělení a analogicky počet narozených kluků do prvního narození holky (b).

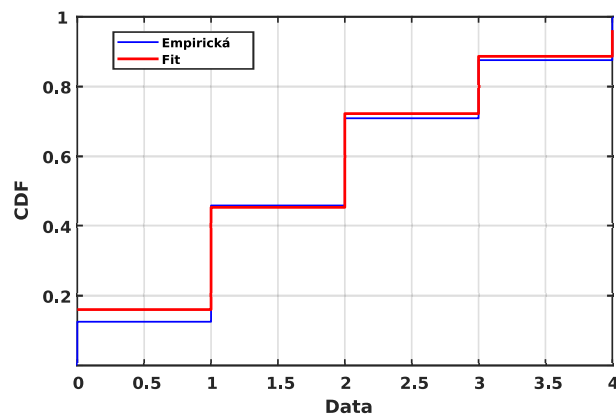
I přes to, že Lillieforsův test zamítnul normalitu dat pro kategorii holek, předpokládejme, že porodní váha kluků i holek mají normální rozdělení a jsou nezávislé¹ a otestujme rovnost rozptylů pomocí dvouvýběrového F-testu. S p-hodnotou 0,075 hypotézu rovnosti rozptylů nezamítáme a následně použijeme dvouvýběrový t-test pro testování rovnosti středních hodnot. Pokud uvažujeme na základě předchozího F-testu, že si rozptyly jsou rovny, dostáváme p-hodnotu 0,135. Pokud uvažujeme rozdílné rozptyly, dostáváme p-hodnotu 0,167. V obou případech hypotézu rovnosti středních hodnot nezamítáme. Pokud použijeme asymptotický t-test, kde předpokládáme obecný i.i.d. \mathcal{L}^2 model, hypotézu opět nezamítáme, tentokrát s p-hodnotou 0,076.

Můžeme také testovat, zda závisí porodní hmotnost na době porodu dopoledne nebo porodu odpoledne. Nejdříve jsme testovali normalitu porodní hmotnosti dětí narozených před dvanáctou hodinou pomocí Pearsonova a Lillieforsova testu se stejným nastavením jako v předchozím odstavci. Pearsonův test sice hypotézu nezamítnul s p-hodnotou 0,053, avšak Lillieforsův test již hypotézu zamítnul s p-hodnotou 0,01. Proto byla normalita hmotnosti dětí narozených před dvanáctou hodinou vyvrácena. Použijeme proto opět asymptotický t-test. Hypotéza nebyla zamítnuta, protože p-hodnota je přibližně rovna 0,258.

¹Případ, kdy by porodní váhy byly závislé, může být například narození vícera. Takové data však nemáme k dispozici.



Obrázek 6: Grafický znázornění dobré shody empirické distribuční funkce času mezi porody (modře) a distribuční funkce exponenciálního rozdělení se střední hodnotou (MLE) 36,6 minut (oranžově).



Obrázek 7: Distribuční funkce příslušné počtu narozených dětí během hodiny (modře) a Poissonově rozdělení se střední hodnotou $\lambda = 1,8333$.

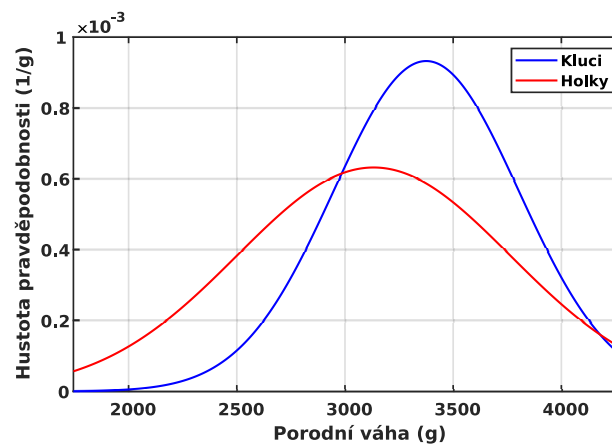
Závěr

Pomocí Pearsonovo χ^2 testu jsme na hladině významnosti 5% ověřili, že počet narozených holek do prvního narození kluka má geometrické rozdělení s parametrem p , který je roven MLE odhadu pravděpodobnosti narození kluka, který je v případě zpracovaných dat přibližně 0,591. Obdobně pak počet narozených kluků do prvního narození holky má geometrické rozdělení s parametrem $1 - p$.

Doba mezi porody má exponenciální rozdělení se střední hodnotou čekání 36,6 minut a počet novorozenců za hodinu pro každou hodinu má Poissonovo rozdělení se střední hodnotou 1,833 novorozenců.

Ikdyž jsme byli neúspěšní v ověření normality porodní váhy holek, ověřili jsme pomocí t-testu na hladině významnosti 5%, že není statisticky významný rozdíl mezi porodní váhou různých pohlaví. Tento výsledek jsme poté potvrdili i legitimním asymptotickým t-testem. Avšak tento výsledek odporuje intuitivní představě, že se kluci rodí těžší, a proto může být závěr zkreslen nízkým počtem provedených měření.

Pomocí asymptotického t-testu jsme na závěr ověřili, že není statisticky významný rozdíl dokonce ani mezi porodní váhou dětí narozených dopoledne a odpoledne.



Obrázek 8: Hustota normálního rozdělení pro porodní váhu kluků (modře) a porodní váhu holek (červeně).