Newborn’s weight data analysis

Daniel Wohlrath

Summer 2020*

Data description

Forty-four children – a new record – were born within twenty-four hours on December 18, 1997, at Mater Mothers’ Hospital in Brisbane, Australia. For each child, data on the time of birth, gender, and birth weight were recorded. The data was published in the Brisbane Sunday newspaper ”The Sunday Mail”on December 21, 1997.

1 Descriptive statistics

Table 1 contains the values of basic statistics for the birth weight of children depending on gender and time of birth. From the listed average values of all children, boys and girls born throughout the day (i.e., the first to third line), we can observe that boys have, on average, a higher birth weight than girls. Moreover, when comparing the respective averages with medians and standard deviations, it is evident that the birth weight of girls has a greater variance, or there were more outliers during measurement that influenced these data.

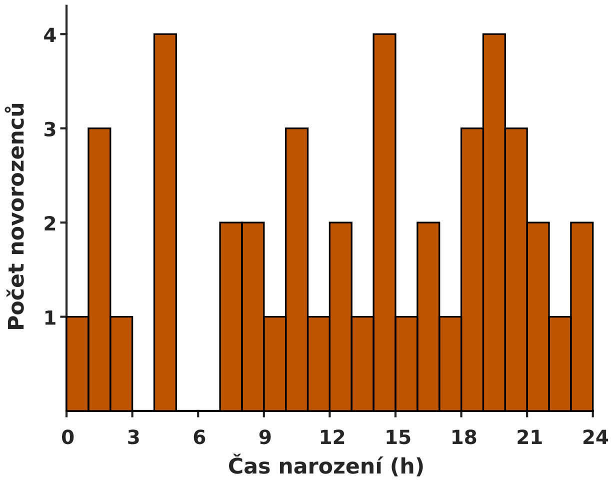
	average (g)	**standard deviation (g)**	**median (g)**	**min (g)**	**max (g)**	**count**
all children	3275.95	528.03	3404	1745	4162	44
boys	3375.31	428.05	3404	2121	4162	26
girls	3132.44	631.58	3381	1745	3866	18
boys (0-12h)	3364.60	316.93	3450	2846	3838	10
boys (12-24h)	3382.00	494.96	3404	2121	4162	16
girls (0-12h)	3022.13	761.91	3271	1745	3837	8
girls (12-24h)	3220.70	531.50	3429	2184	3866	10
all children (0-12h)	3212.39	568.22	3357	1745	3838	18
all children (12-24h)	3319.96	505.06	3417	2121	4162	26

Tabulka 1: Table showing basic statistical data for the measured data, divided into categories by gender and time of birth.

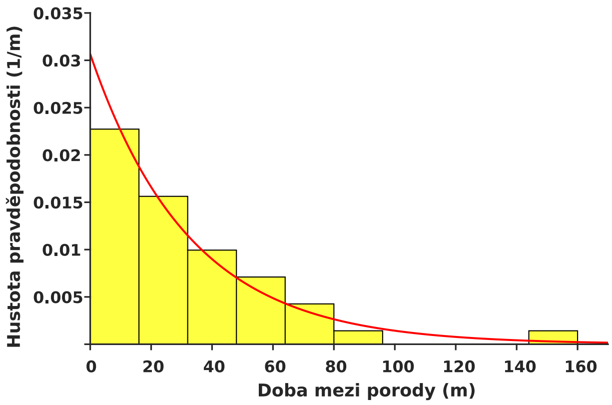
*This report was translated into English only later

Figures 1 and 2 present histograms showing the number of births in each hour of the day and the time intervals between individual births. The average time between two births is 32 minutes and 36 seconds. Furthermore, Figure 2 clearly shows that the data very well correspond to an exponential distribution, a fact that will be verified in the next part of this text.

The occurrence of childbirth may or may not be independent of other births. Since, according to the source, the published data was record-breaking, it is possible that midwives and obstetricians were very busy on this day. Therefore, it might not have been possible for more births to occur in a short time than the number of available doctors, and thus births could have been influenced by various drugs, for example, to delay them until the doctors arrive from another birth. Since we do not have such data available, let us assume in the further analysis that such influencing of births did not occur, and therefore the individual birth times are independent events.

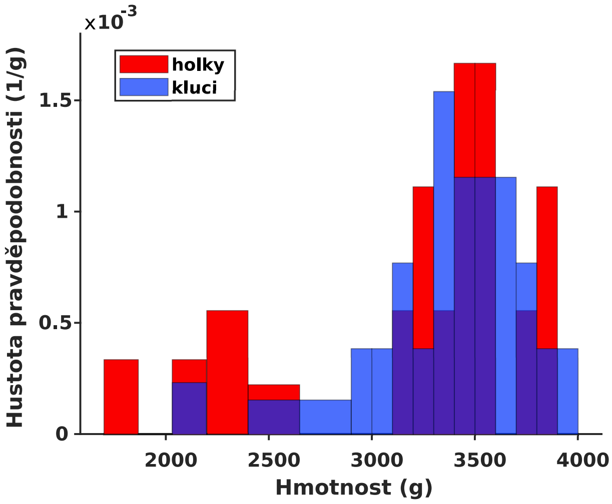


Obrázek 1: Frequency of newborns born in each hour



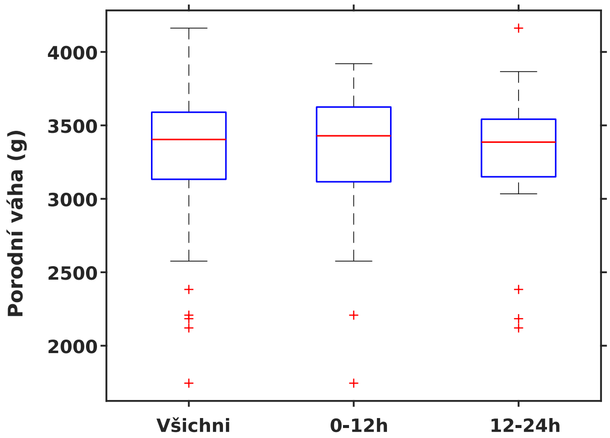
Obrázek 2: Histogram estimate of probability density (yellow columns) and probability density of exponential distribution with mean equal to MLE (red solid line) for the waiting time before the next birth.

From Figure 3, it is evident that girls are born at lower birth weights more frequently than boys. This fact can be mathematically interpreted as the left tail of the probability density of birth weight for boys decreases more rapidly than the left tail for girls. The boxplot in Figure 4a corresponds to the notion that the birth weight of newborns is probably not dependent on the time of birth. On the other hand, the boxplot in Figure 4b points to the fact that the average

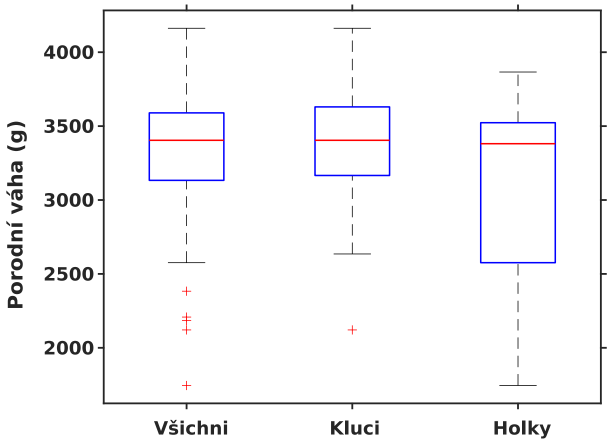


Obrázek 3: Histogram estimate of probability density for the weight of newborn boys (blue) and girls (red).

birth weight of girls is lower than that of boys, while the medians of the individual categories are approximately the same.



(a) Boxplot for the birth weight of children born in the morning, afternoon, and throughout the day.



(b) Boxplot for the birth weight of boys, girls and both genders simultaneously.

Obrázek 4: Boxplot for the birth weight of children of different categories.

2 Custom Analysis

For all the following tests, a significance level of $\alpha = 0.05$ was chosen. In the Pearson χ^2 goodness-of-fit tests mentioned, we combined adjacent cells if the number of observations in individual cells was less than 5. If a different merging rule was applied, it is explicitly mentioned. The calculations were carried out in the MATLAB programming environment.

2.1 Goodness-of-Fit Tests

Refer to Table 1, out of a total of 44 born children, 26 are boys and 18 are girls. Therefore, the Maximum Likelihood Estimate (MLE) for the probability of a boy being born is 0.591, with a 95% confidence interval of $(0.433, 0.737)$.

First, let's test the hypothesis whether the number of born girls or boys until the first boy or girl has a geometric distribution. In the first case, we test the hypothesis

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

where F is the distribution function corresponding to the mentioned random variable, and p is the probability of a boy being born (success) in the geometric model. The MLE estimate for the parameter p is 0.5. We divided the range of the variable into 4 bins centered at points $0, 1, 2, 3$. Using the Pearson χ^2 test, we obtained a p-value of 0.389 and a Pearson statistic value of 1.889, which is less than $\chi^2_{0.95}(2)$. Therefore, we do not reject the hypothesis. A comparison of observed and expected frequencies in individual bins is provided in Table 2.

Value	0	1	2	3
Observed count	3	4	1	1
Expected count	4,50	2,25	1,13	1,13

Tabulka 2: Table depicting empirical and expected frequency of observations of the number of born girls until the first boy is born.

Similarly, we tested the hypothesis using Pearson's χ^2 test for the second variant, where the Maximum Likelihood Estimate (MLE) of the relevant parameter is 0.308. With the same significance level and division of the domain into 5 bins centered at points $0, 1, 2, 3, 4$, we obtained a p-value of 0.252, and therefore, we do not reject the hypothesis at this time.

The graphical representation of the corresponding distribution functions is shown in Figures 5a and 5b, and the comparison of observed and expected frequencies in individual bins is provided in Table 3.

Value	0	1	2	3	4
Observed count	2	0	2	2	2
Expected count	2,46	1,70	1,18	0,82	1,84

Tabulka 3: Table depicting empirical and expected frequency of observations of the number of born boys until the first girl is born.

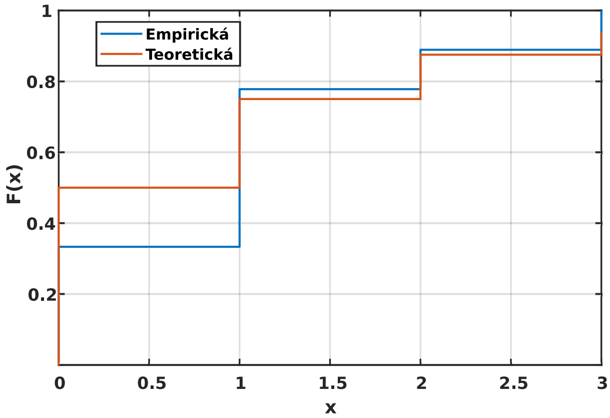
Let's further test the hypothesis whether the inter-birth interval follows an exponential distribution. The graphical visualization of this test is depicted in Figures 2 and 6. We initially conducted the Pearson's χ^2 test, which did not reject the hypothesis with a p-value of 0.851, indicating very good agreement between the data and the theoretical model. Additionally, we used the Lilliefors test, which also did not reject the null hypothesis, with a p-value of 0.247.

Let's consider the hypothesis that the number of newborns per hour for each hour follows a Poisson distribution. The frequency of observations of this variable is shown in Table 4. The estimated value of the parameter is $\lambda = 1.833$, and we verified the agreement between the data and the model through a graphical comparison of the distribution functions (see Figure 7) and the Pearson's χ^2 test, which did not reject the hypothesis with a p-value of 0.958.

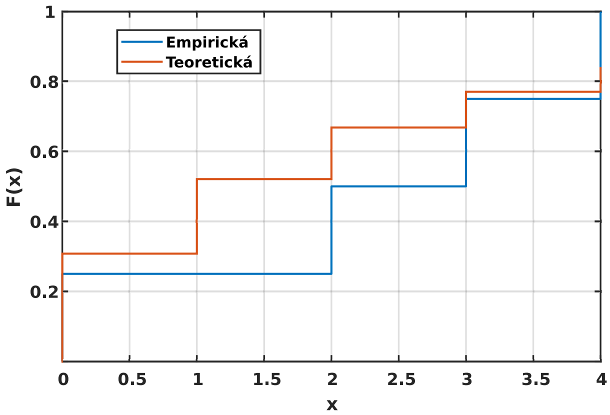
2.2 Hypothesis Testing

The main question is whether birth weight depends on gender. First, let's test whether the distribution of birth weights for boys and girls is Gaussian.

Using the Pearson's test to evaluate the hypothesis that the birth weight of boys follows a normal distribution with estimated parameters $\mu = 3375.31$ and $\sigma = 428.05$, we obtained



(a)



(b)

Obrázek 5: Comparison of empirical and theoretical distribution functions when testing the hypothesis whether the number of born girls until the first boy is born (a) follows a geometric distribution, and analogously, the number of born boys until the first girl is born (b).

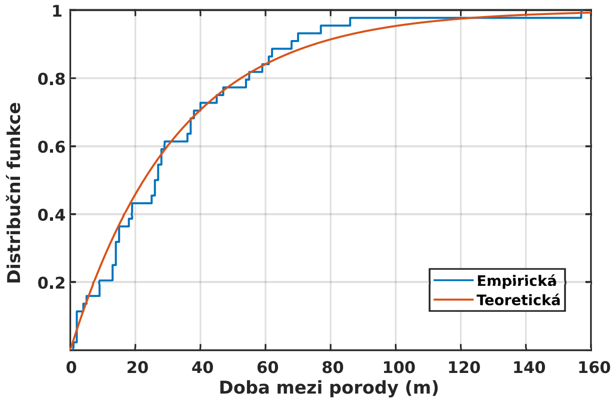
a p-value of 0.42. We adjusted the merging rule for bins when there were a small number of observations, combining them into pairs. The hypothesis was not rejected, even when using the Lilliefors test, with a p-value of 0.103.

Similarly, we proceeded to test the hypothesis for girls, where $\mu = 3132.44$ and $\sigma = 631.58$. We applied Pearson's test, again with the adjusted merging rule (combining two observations), resulting in a p-value of 0.135. Therefore, this test did not reject the hypothesis. However, the Lilliefors test rejected the normality hypothesis with a p-value of 0.028. Subsequently, another Lilliefors test was performed, considering the null hypothesis that the distribution of birth weights for girls follows a Weibull distribution with parameters $A = 3371.18$ and $B = 6.86$. The hypothesis was not rejected, as the p-value was 0.137. Similarly, Pearson's test did not reject the hypothesis of a Weibull distribution with a p-value of 0.274. A comparison of the estimated probability densities of the normal distribution is shown in Figure 8.

2.3 Hypothesis Testing

Despite the Lilliefors test rejecting normality for the category of girls, let's assume that the birth weights of both boys and girls follow a normal distribution and are independent[1], and test the

[1]The case where birth weights are dependent could be, for example, the birth of multiples. However, such data is not available.



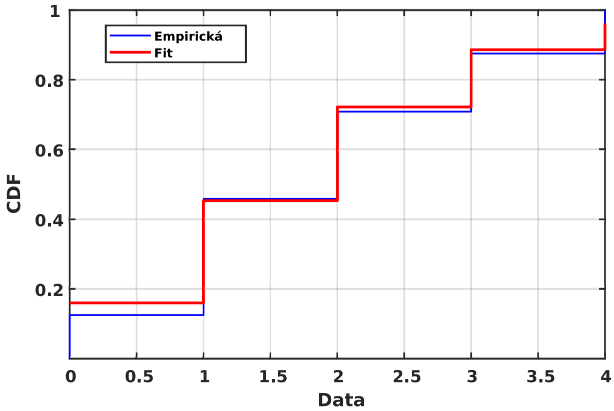
Obrázek 6: Graphical representation of the goodness-of-fit between the empirical distribution function of the inter-birth interval (in blue) and the distribution function of the exponential distribution with a mean (MLE) of 36.6 minutes (in orange).

Newborn per hour	0	1	2	3	4
Observed count	3	8	6	4	3
Expected count	3,84	7,03	6,45	3,94	2,74

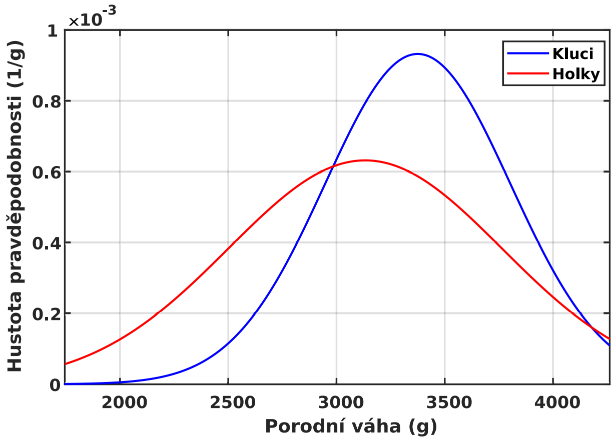
Tabulka 4: Table of observed and expected frequencies of newborns per hour (e.g., During measurements, the birth of exactly one child was observed eight times within one hour).

equality of variances using a two-sample F-test. With a p-value of 0.075, we do not reject the hypothesis of equal variances. Subsequently, we use a two-sample t-test to test the equality of means. Considering the previous F-test and assuming equal variances, we obtain a p-value of 0.135. Assuming different variances, we get a p-value of 0.167. In both cases, we do not reject the hypothesis of equal means. If we use the asymptotic t-test, assuming a general i.i.d. $\mathcal{L}^2$ model, the hypothesis is again not rejected, this time with a p-value of 0.076. Furthermore, without assuming normality, we can use the Wilcoxon test, which, similar to the asymptotic t-test, does not assume a normal distribution. The hypothesis that the variables come from a continuous distribution with the same median is not rejected by the Wilcoxon test, with a p-value of 0.352.

We can also test whether birth weight depends on the time of birth, either in the morning or in the afternoon. First, we tested the normality of the birth weight of children born before noon using Pearson's and Lilliefors' tests with the same settings as in the previous paragraph. Although the Pearson test did not reject the hypothesis with a p-value of 0.053, the Lilliefors test already rejected the hypothesis with a p-value of 0.01. Therefore, the normality of the birth weight of children born before noon was disproven. We then used the asymptotic t-test again. The hypothesis was not rejected because the p-value is approximately 0.258. The Wilcoxon test was also used, which did not reject the hypothesis with a p-value of 0.659, that the data comes from a distribution with the same median.



Obrázek 7: Distribution function of the corresponding number of births during an hour (in blue) and a Poisson distribution with a mean of $\lambda = 1.8333$.



Obrázek 8: Probability density of the normal distribution for the birth weight of boys (in blue) and the birth weight of girls (in red).

Conclusion

Using the Pearson's χ^2 test at a significance level of 5%, we verified that the number of born girls until the first boy has a geometric distribution with parameter p, equal to the Maximum Likelihood Estimate (MLE) of the probability of a boy being born, which is 0.5 in the processed data. Similarly, the number of born boys until the first girl has a geometric distribution with a parameter of 0.308.

The inter-birth interval follows an exponential distribution with a mean waiting time of 36.6 minutes, and the number of newborns per hour for each hour follows a Poisson distribution with a mean of 1.833 newborns.

Although we were unsuccessful in verifying the normality of the birth weight of girls, we used a t-test at a significance level of 5% to confirm that there is no statistically significant difference in birth weight between different genders. This result was then confirmed by a legitimate asymptotic t-test and the Wilcoxon test. However, this result contradicts the intuitive notion that boys are born heavier, possibly distorted by the low number of measurements conducted.

Finally, using the asymptotic t-test and the Wilcoxon test, we verified that there is no statistically significant difference even in the birth weight of children born in the morning and afternoon.