STATISTICAL INFERENCE (MS-C1620). 8.1.–17.4.2025. Aalto University. University Lecturer Pekka Pere.

Suggested solutions to the questions of exam 5.6.2024

1.

a) The sample size (n) is fairly large $2548 + 132 = 2680$, and apparently the estimated proportion (π) is not very small. The usual 99% Wald confidence interval

$$\left[\hat{\pi} - 2.575829\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + 2.575829\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right]$$

should be reliable (`qnorm(0.995)` returns the 0.995th quantile of the Standard Normal distribution). The upper and lower bounds of the 99% confidence interval for the proportion of biological mothers who commit severe violent acts toward their child are

$$\begin{aligned} 0.04925373 \pm 2.575829 \times \sqrt{\frac{0.04925373(1-0.04925373)}{2680}} \\ = 0.04925373 \pm 0.01076717 \end{aligned}$$

($132/2680 = 0.04925373$). The confidence interval is

$$[0.038, 0.060].$$

Let us check our calculation with command `binom.test` of the mosaic package:

```
install.packages("mosaic") # only when downloading package 1st time
library(mosaic)
binom.test(x=132,n=2680,conf.level=0.99,ci.method="Wald")
 - -
## data:  132 out of 2680
## number of successes = 132, number of trials = 2680, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.5
## 99 percent confidence interval:
##  0.03848656 0.06002090
## sample estimates:
## probability of success
##             0.04925373
```

Command `binom.test` returns the same confidence interval.

b) There are only 28 observations of step/adoptive/foster mothers. Agresti–Caffo is a more reliable confidence interval than the Wald interval if the sample size is small. Let us hence calculate the Agresti–Caffo confidence interval. The calculation starts by adding a pseudo observation to each cell of the original table:

	severe violence		
	yes	no	Σ
biological mother	133	2549	2682
step/adoptive/foster mother	10	20	30

The estimates of the proportions are $10/30 = 1/3$ and $133/2682 = 0.04958986$. Let us calculate the conventional 99% Wald confidence interval from the adjusted data:

$$\left[\hat{\pi}_1 - \hat{\pi}_2 - 2.575829\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}},\right.$$
$$\left.\hat{\pi}_1 - \hat{\pi}_2 + 2.575829\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}\right]$$

The square-root term from the adjusted data is

$$\sqrt{\frac{(1/3)\times(1-1/3)}{30} + \frac{0.04958986\times(1-0.04958986)}{2682}} = 0.08616833.$$

The upper and lower bounds of the 99% confidence interval are

$$1/3 - 0.04958986 \pm 2.576 \times 0.08616833 = 0.2837435 \pm 0.2219696.$$

Agresti–Caffo 99% confidence interval is

$$[0.062, 0.506].$$

A check of the correctness of our calculations is obtained with the `wald2ci` command of the PropCIs package:

```
install.packages("PropCIs") # only when downloading package 1st time
library(PropCIs)
wald2ci(9,28,132,2680,adjust="AC")
 - -
## 99 percent confidence interval:
##  0.06178858 0.50569837
## sample estimates:
## [1] 0.2837435
```

The same confidence interval was obtained.

The confidence interval does not cover zero. The inference is that in the population the proportion of step/adoptive/foster mothers committing severe violent acts toward their child is larger than the corresponding proportion of biological mothers.

Let us calculate the corresponding 99% Wald confidence interval for a comparison:

```
wald2ci(9,28,132,2680,conf.level=0.99,adjust="Wald"))
 - -
## 99 percent confidence interval:
##  0.04457884 0.49977084
##
## [1] 0.2721748
```

The Wald confidence interval $(0.046, 0.500)$ is shifted to the left in comparison to the Agresti–Caffo confidence interval $(0.062, 0.506)$.

The main finding of the study was that 5.5% of mothers had committed severe violent acts toward their (biological, step *etc.*) child at least once during the preceding 12 months. In a kindred article the corresponding figure for fathers is 5.9%.[1]

[1] N. Ellonen, K. Peltonen, T. Pösö ja S. Janson (2017): A Multifaceted Risk Analysis of Fathers' Self-Reported Physical Violence Toward Their Children. *Aggressive Behavior*, 43, 317–328.

2.

a) The sign-test statistic is

$$\frac{S/n - 0.5}{\sqrt{0.5(1-0.5)/n}}.$$

Here S is the number of pluses or in the present context the number of positive returns. The sample size is n.

b) The numerical value of the test statistic is 6.621:

$$\frac{771/1303 - 0.5}{\sqrt{0.5(1-0.5)/1303}} = 6.621.$$

c) The statistic follows the standard Normal distribution for large samples as the one here. The critical value from the Standard Normal distribution at the 0.001 significance level is $3.290 < 6.621$ (two-sided test). The null hypothesis is rejected. The conclusion is that the median return is larger than zero.

3.

a) Interpretation:

- If economic losses are 0 units then the expected value of fine in the same units is the value of intercept or 2441.80.
- If economic losses x increase by a unit then the expected value of fine Y increases by 40.3 units.
- The coefficient of determination is

$$\mathsf{R}^2 = 1 - \frac{\mathsf{RSS}}{\mathsf{TSS}} = \frac{\mathsf{RSS} - \mathsf{TSS}}{\mathsf{TSS}}.$$

  Above $\mathsf{RSS} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is the residual sum of squares ($\hat{y}_i$ is the fit for the ith observation), and $\mathsf{TSS} = \sum_{i=1}^{n}(y_i - \bar{y})^2$ is total sum of squares ($\bar{y}$ is the sample mean of y_is). Thus the coefficient of determination is the portion of the sample variance of y which the model can explain. For the present model the coefficient of determination is 0.182, i.e. the model explains 18.2% of the variance of the fine.

b) In the case of a single regressor only it holds that $\mathsf{R}^2 = \hat{\rho}^2$ or $\hat{\rho} = \pm\sqrt{\mathsf{R}}$, where $\hat{\rho}$ is sample correlation. For the present model

$$\hat{\rho} = \pm\sqrt{0.182} \approx \pm 0.427.$$

It is clear from the figure that the sample correlation is positive so $\hat{\rho} \approx 0.427$. The sample correlation between illegitimate strikes and economic losses is 0.427.

The government of Finland is presently (2024) planning to raise fines from illegitimate strikes.

c) Such outstanding observations are called outliers. One should not delete outliers automatically from the data. Outliers may sometimes result from random variation in accordance with the assumed model. In such cases deletion would likely to be inappropriate. If the outlier is due to a recording error, an exceptional circumstance outside the scope of the model *etc.* then deletion would be in order.

4.

a) The group means are in accordance with the theory of the researchers. Hebrew reading test subjects read from the right to the left. Calorie content is smallest on average with the group who got calorie information to the right of the course. Also interestingly, average calorie content is greatest with the group which did not get any calorie information.

b) The degrees of freedom of the F-statistic are in general $m-1$ and $n-m$, where m is the number of groups and n is the sample size. In the question $m=3$ and $n=85+86+81=252$. The degrees of freedom are hence $3-1=2$ and $252-3=249$.

c) The p-value of the F-statistic 0.06 is larger than the significance level of the test. The null hypothesis is hence not rejected. The test suggests that the null hypothesis of no difference between the group means should not be rejected — despite that the numerical values of the group means follow the theory.