STATISTICAL INFERENCE (MS-C1620). 8.1.–17.4.2024. Aalto University. University Lecturer Pekka Pere.

Suggested solutions to the questions of exam 27.8.2024

1. No parent out of 325 responded that his or her child does not own a phone. The rule of three gives [0,0.01] as the 95% confidence interval for the proportion of children without a phone:

$$\left[0, \frac{3}{325}\right] \approx [0, 0.01].$$

2.

a) The sign test statistic and the value of it:

$$\frac{S/n - 0.5}{\sqrt{0.5(1-0.5)/n}} = \frac{82/82 - 0.5}{\sqrt{0.5(1-0.5)/82}} = 9.055385.$$

Under the null hypothesis the statistic follows approximately the Standard normal distribution by the central limit theorem. The 0.999th quantile of the Standard normal distribution is 3.090 (`qnorm(0.999)`). The null hypothesis of equal probability of a newborn boy or girl is rejected at significance level 0.002 because $9.055 > 3.090$ (two-sided test).

R code below returns the value of the sign test statistic and its p-value. The ratio figure of the number of newborn boys and girls is in data frame `Arbuthnot` under name `Ratio`. The code lower down (command `SIGN.test(Arbuthnot$Ratio)`) returns statistics related to the sign test statistic:

```
# Lasketaan merkkitestisuure ja sen p-arvo standardinormaalijakaumasta:
s <- 82
n <- 82
(s/n-0.5)/sqrt(0.5*(1-0.5)/n)
## [1] 9.055385
2*pnorm(-9.055385)
## [1] 1.360869e-19

install.packages("BSDA") # vain 1. kerran pakettia käytettäessä
library(BSDA) # package Basic Statistics and Data Analysis
SIGN.test(Arbuthnot$Ratio)
##         One-sample Sign-Test
## data:  Arbuthnot$Ratio
## s = 82, p-value = 2.22e-16
## alternative hypothesis: true median is not equal to 0
## 95 percent confidence interval:
##  1.059203 1.076478
## sample estimates:
## median of x
##    1.064704
##  - -
```

b) Command `shapiro.test(Arbuthnot$Ratio)` has calculated the value of the sign test statistic 0.960 and its p-value 0.011:

```
install.packages("BSDA") # vain 1. kerran pakettia käytettäessä
library(BSDA) # package Basic Statistics and Data Analysis
shapiro.test(Arbuthnot$Ratio)
##         Shapiro-Wilk normality test
## data:  Arbuthnot$Ratio
## W = 0.95956, p-value = 0.01117
```

The null hypothesis of the test is that the data has originated from a Normal distribution. The null hypothesis is not rejected at significance level 0.01 because $0.011 > 0.01$ or because the p-value of the test statistic is larger than the significance level of the test.

3.

a) The sample variance of the charring rate is almost 60-fold in the 2nd layer ($0.1669577/0.002843024 \approx 58.7$) and the sample standard deviation is almost 8-fold ($(0.1669577/0.002843024)^{1/2} \approx 7.7$)! The sample variances are very likely to differ statistically significantly. The test statistic is simply the ratio of the sample variances:

$$\frac{s_1^2}{s_2^2} = \frac{0.002843024}{0.1669577} = 0.01702841.$$

The test statistic follows the $\mathsf{F}(52, 52)$-distribution under the null hypothesis ($\mathsf{F}(n_1 - 1, n_2 - 1)$-distribution for sample sizes n_1 and n_2).

The p-value is calculated from the left tail of the distribution because the value of the test statistic is smaller than 1 and the alternative hypothesis is $\sigma_1^2 < \sigma_2^2$. The p-value is here the probability that the test statistic takes value 0.01702841 or smaller under the null distribution $\mathsf{F}(52, 52)$. (The probability is zero to 32 decimal places according to output `1.090972e-32` from command `pf(0.01702841,52,52)`.)

If one takes

$$\frac{s_2^2}{s_1^2} = \frac{0.1669577}{0.002843024} = 58.72539,$$

as the test statistic then the right-tail of the null distribution is at focus. Now the p-value is 1 minus the probability to the right of quantile 58.72539 of the $\mathsf{F}(52, 52)$ distribution. (Command `1-pf(58.72539,52,52)` returns 0. It agrees with the previous calculation of the p-value within numerical accuracy.)

Let us check with R that the test statistic and the p-value of it have been calculated correctly and that R agrees with our reasoning of the degrees of freedom. The appropriate command is `var.test` (the observations from the 1st and 2nd layers are in variables `kerros1` and `kerros2` of dataframe `datai`):

```
var.test(datai$kerros1,datai$kerros2,alternative="less")
##          F test to compare two variances
## data:  datai$kerros1 and datai$kerros2
## F = 0.017028, num df = 52, denom df = 52, p-value < 2.2e-16
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 0.026985
## sample estimates:
## ratio of variances
##         0.01702841
```

R calculates the same value of the test statistic (and p-value of it within calculation or reporting accuracy) and returns the same degrees of freedom as stated above.)

b) The test statistic and its value are

$$\frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{0.6818273 - 1.116031}{\sqrt{\frac{0.002843024}{53} + \frac{0.1669577}{53}}} = -7.671171.$$

The null distribution of the statistic is t. The degrees of freedom of it ν are calculated with (the Smith–Welch–Satterthwaite-) formula:

$$\begin{aligned}
\nu &= \mathsf{int}\left[\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}\right] \\
&= \mathsf{int}\left[\frac{\left(\frac{0.002843024}{53}+\frac{0.1669577}{53}\right)^2}{\frac{\left(\frac{0.002843024}{53}\right)^2}{52}+\frac{\left(\frac{0.1669577}{53}\right)^2}{52}}\right] \\
&= \mathsf{int}\,[53.77044] \\
&= 53.
\end{aligned}$$

The approximate null distribution of the test statistic is $\mathtt{t}(53)$. The sample size or the degrees of freedom are large enough so that the corresponding t-distribution should agree fairly well with the Standard normal distribution. The Standard normal can hence also be used as an approximative null distribution. The 0.0005th quantiles of the $\mathtt{t}(53)$ and Standard normal distributions are -3.484 (`qt(0.0005,53)`) and -3.291 (`qnorm(0.0005)`), respectively. The value of the test statistic is larger than either of the fractiles. The null hypothesis is rejected at significance level 0.001. Charring rate is larger in the 2nd layer than in the 1st layer.

A check with command `t.test`:

```
t.test(datai$kerros1,datai$kerros2,var.equal=FALSE)
##          Welch Two Sample t-test
## data:  datai$kerros1 and datai$kerros2
## t = -7.6712, df = 53.77, p-value = 3.443e-10
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5476953 -0.3207127
## sample estimates:
## mean of x mean of y
## 0.6818273 1.1160313
```

R calculates the same value of the test statistic and reports the same degrees of freedom as above.

4.

a) The different variances of the charring rates in layers 1 and 2 are not a barrier for applying the pairwise t test. The test assumes that the variance of the pairwise differences is constant. Equality of variances in layers 1 and 2 need not hold.

b) The sample mean and variance of the pairwise differences of the charring rates are -0.434204 and 0.1789338, respectively. The test statistic is

$$\frac{\hat{\mu}_D}{s_D/\sqrt{n}} = \frac{-0.434204}{\sqrt{0.1789338}/\sqrt{53}} = -7.472838.$$

The null distribution of the statistic is $\mathtt{t}(52)$($\mathtt{t}(n-1)$ in general). The 0.0005th quantile of it is -3.487691 (`qt(0.0005,52)`). The sample size or the degrees of freedom are large enough so that the null distribution can be approximated by the Standard normal distribution. The 0.0005th quantile of it is -3.290527 (`qnorm(0.0005)`). The value of the test statistic-7.472838 is much smaller than either of the quantiles. The null hypothesis is rejected at significance level 0.001. Charring rate is larger in the 2nd layer than in the 1st layer.

A check with command `t.test` now with argument `paired=TRUE`:

```
t.test(datai$kerros1,datai$kerros2i,paired=TRUE)
##         Paired t-test
## data:  datai$kerros1 and datai$kerros2i
## t = -7.4728, df = 52, p-value = 8.73e-10
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  -0.5507988 -0.3176092
## sample estimates:
## mean difference
##       -0.434204
```

R calculates the same value of the test statistic and reports the same degrees of freedom for the `t` distribution as above.