

Task 1

Reading papers: List 5 important statements

“Do baseline p-values follow a uniform distribution in Randomised Trials?”

- 1) It depends on the data kind (e. g. binary data) and distribution (e. g. log-normal distribution), way of testing (e.g. t-test) and independence of data points if the p-values come close to a uniform distribution
- 2) For normal distribution with independent data points and used t-test we get uniform distribution of p-values when using randomized trials
- 3) No uniform distribution for p-values when using: skewed data distribution; correlated two sample t-test; large samples with chi-squared test
- 4) Idea of uniform distribution of p-values in many circumstance not accurate. Therefore, danger of wrong conclusions in research given
- 5) criticism about p-values in literature: When both groups of comparison are chosen randomly, there are drawn from the same population and therefore null-hypothesis shall become true.

“P-values are random variables”

- 1) Assumption: Students are wrongly taught about the meaning of p-values
- 2) p-value is no probability of the null-hypothesis being true
- 3) Test of independence between categorical data: use chi-square test or Fisher’s exact test. However, we end up with a discrete distribution of p-values and histograms of p-values without any use. Not all p-values are achievable when using discrete data.
- 4) Possibility of exploring accuracy of asymptotic approximations in different tests
- 5) Even with only changing the mean of a distribution the p-value distribution can modify.

(Task 2 next page)

Task 2

Testing of composite hypothesis of normality of residuals by using five omnibus tests from R package nortest

In general (not discussed so far):

- an alternative hypothesis is always a composite hypothesis
- omnibus test tests for significance of several parameters in a model at once. e. g. When null-hypothesis say $\mu_1 = \mu_2 = \mu_3$ (all population means equal) → here we have 3 parameters at once. For instance, the test appears in ANOVA models or multiple linear regression models.
- Documentation “nortest”: <https://cran.r-project.org/web/packages/nortest/nortest.pdf>

Needed data:

```
birds <- c( 57,31,25,32,33,10,18,9,12,14,10,26,4,16,13)
```

```
altitude <- c(2397,1264,1170,1060,1900,460,1998,700,735,613,662,1502,750,2275,552)
```

```
lin_model <- lm(altitude~birds)
```

```
residuals_ <- lin_model$residuals
```

We shall use this package to check out the residuals:

```
#install.packages("nortest") # install the needed package # here commented out
```

```
library(nortest) # load package
```

1. Anderson_Darling test for normality:

```
p1 <- ad.test(residuals_)$p.value
```

2. Cramer-von Mises test for normality:

```
p2 <- cvm.test(residuals_)$p.value
```

3. Lilliefors (Kolmogorov-Smirnov) test for normality:

```
p3 <- lillie.test(residuals_)$p.value
```

4. Pearson chi-square test for normality:

```
p4 <- pearson.test(residuals_, n.classes=length(residuals_))$p.value
```

5. Shapiro_Francia test for normality:

```
p5 <- sf.test(residuals_)$p.value
```

```
p_average <- (p1+p2+p3+p4+p5)/5
```

```
> p1
[1] 0.01000786
> p2
[1] 0.01455709
> p3
[1] 0.113971
> p4
[1] 0.3007083
> p5
[1] 0.009143305
> p_average
[1] 0.08967751
```

The tests refer to the composite hypothesis of normality of the inputted x (here residuals). $p_average = 0.08967751 > significance_level = 0.05$ which means that we **can't** reject the null-hypothesis that the residuals come from a normal distribution. We can integrate that knowledge when creating a second better fitting model.

(Task 3 next page)

Task 3

Petrol prices

A normal distribution may help us to recognize unusual price changes (when the residual becomes to large).

Used in lecture for sequential data: Shapiro-Wilk normality test

```
# Inspired by https://at.fuelo.net/?lang=en:
time <- c(1,2,3,4,5,6,7)
price <- c(1.695, 1.69,1.675,1.68,1.685,1.72,1.69)

petrol_model <- lm(price~time)

petrol_residuals <- petrol_model$residuals

shapiro_wilk_test <- shapiro.test(petrol_residuals)
```

```
> shapiro_wilk_test <- shapiro.test(petrol_residuals)
> View(shapiro_wilk_test)
> shapiro_wilk_test

      Shapiro-Wilk normality test

data:  petrol_residuals
W = 0.89322, p-value = 0.2919
```

For the Shapiro-Wilk normality test as used for sequential data in the lecture we observe p-value > 0.05. Therefore, we can't reject the null-hypothesis, so the residuals seem to be normally distributed.

However, we only have drawn seven observations (at least four required) out of the population. So, more data points could give a better hint whether the residuals are really from a normal distribution.

```
# Let's come to the second part of the task:
price2 <- price[-1] # remove first element from price vector
price3 <- price[-length(price)] # remove last element

jump <- price2-price3

jump_model <- lm(jump~time[-1])
plot(jump_model)

price2 <- price[-1] # remove first element from price vector
price3 <- price[-length(price)] # remove last element

jump <- price2-price3
```

```
jump_model <- lm(jump~time[-1])
```

```
jump_residuals <- jump_model$residuals
```

```
shapiro_wilk_test2 <- shapiro.test(jump_residuals)
```

```
> shapiro_wilk_test2  
  
      Shapiro-Wilk normality test  
  
data:  jump_residuals  
W = 0.96963, p-value = 0.8899
```

p-value is reaching now a much higher score. To concatenate observations/samples of the same type but at different time points helps the model to realize connections. So, also the residuals have stronger connections due to the link of samples. Therefore, it's easier to guess if they are from the normal distribution.