Non-Parametric Tests

Types of Statistical Tests

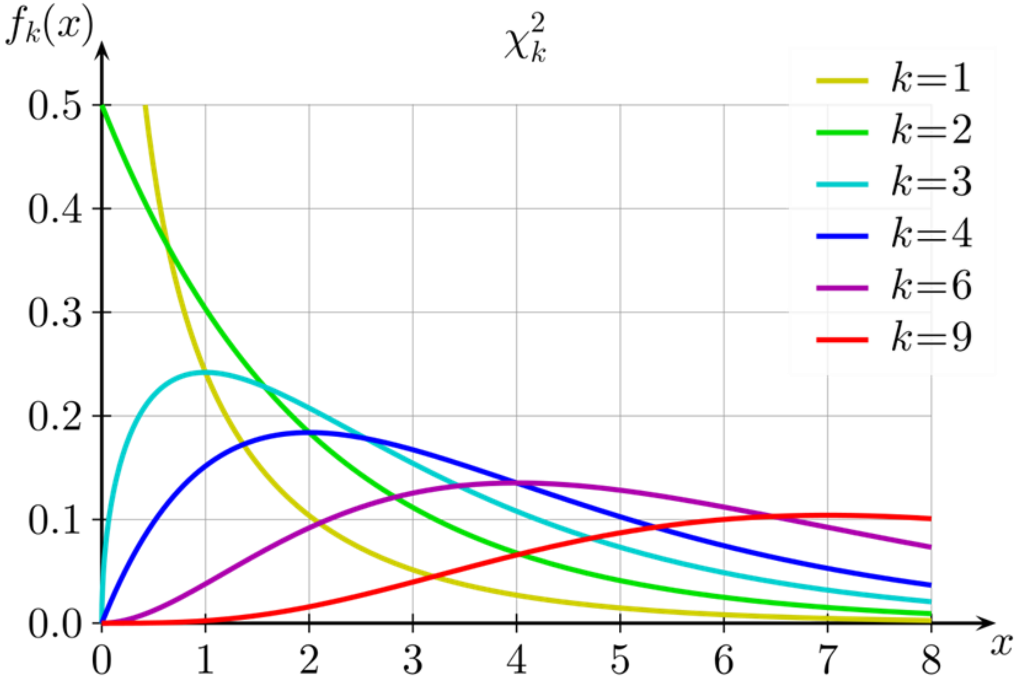
- One-sample parametric tests
  - expected value
  - proportion
  - variance / standard deviation
- Two independent samples parametric tests
  - expected value
  - proportion
  - variance / standard deviation
- Multiple independent samples parametric tests
  - expected value
- Large-sample non-parametric tests
  - goodness-of-fit test
  - test of independence
  - test of homogeneity

LARGE-SAMPLE NON-PARAMETRIC TESTS

- The null hypothesis pertains not to a parameter but to the distribution of a single population or the comparison of two populations through their distributions
  - the population and sample are analyzed by dividing them into bins (categories)
  - these tests are applicable only for large samples
    - each bin should contain at least 10 elements for theoretical (hypothetical) values → or an absolute minimum of 5 elements
  - right-tailed χ^2 tests.
  - the test-name reflects the statement of the null hypothesis.
- Types of Tests:
  - Goodness-of-Fit Test
  - Test of Homogeneity

The χ^2 Distribution – Reminder

- Expected Value = Degrees of Freedom (k)
  $\Rightarrow$ median $\approx$ degrees of freedom
- It's the sum of k squared $N(0,1)$ distributions
- The values corresponding to the extreme 5% are not symmetrically distributed around the median
  - the value corresponding to the lower 5% is closer to the median than the value corresponding to the upper 5% (long right tail)
  - as the degrees of freedom increase, this asymmetry decreases
    (limits to the $N(k, 2k)$ distribution)



Goodness-of-Fit Tests

- It is used to test a hypothesis about the distribution of the population.
  - the population under investigation is divided into k bins based on a certain variable ($C_i \to P(C_i)$), and the same binning is applied to the sample.

Bin (C_i)	Population	Sample	
	P_i	n_i	n_i^*
C_1	P_1	n_1	n_1^*
C_2	P_2	n_2	n_2^*
…	…	…	…
C_k	P_k	n_k	n_k^*
Sum	1	n	n

- Types:
  - pure: distribution parameters are known.
  - estimated: some (b) parameters are estimated from the sample.

Process of Goodness-of-Fit Tests

- assumption
  - large sample size: $\forall i: nP_i \geq 10$
- Test:
  - $H_0: \forall i: P(C_i) = P_i \Leftrightarrow H_1: \exists i: P(C_i) \neq P_i$
    - H_0: the empirical distribution $P(C_i)$ matches a presumed theoretical distribution P_i
  - test statistic: $\chi^2 = \sum_{i=1}^{k} \frac{(n_i - n_i^*)^2}{n_i^*} = n \cdot \left(\sum_{i=1}^{k} \frac{g_i^2}{P_i} - 1 \right)$
  - distribution: right-tailed $\chi^2(k - b - 1)$
  - k = number of bins
  - b = number of estimated parameters

CORVINUS UNIVERSITY of BUDAPEST

Test of Representativeness

We aim to estimate whether a sample of 1000 individuals drawn from the 15–74 age group is representative in terms of highest educational attainment at the 5% significance level.

Education Level	Population Proportion	Sample Size
Less than 8 years of primary school	1.1%	4
Primary school (8 years)	17.5%	160
Vocational school	22.7%	249
Secondary school without vocational training	13.7%	125
Secondary school with vocational training	20.0%	223
College or Bachelor's degree	13.2%	138
University or Master's degree	11.8%	101
Total	**100%**	**1000**

- Sample size is large as $\forall i: n_i^* \geq 10$

Education Level	P_i	n_i	n_i^*
Less than 8 years of primary school	0.011	4	11
Primary school (8 years)	0.175	160	175
Vocational school	0.227	249	227
Secondary school without vocational training	0.137	125	137
Secondary school with vocational training	0.200	223	200
College or Bachelor's degree	0.132	138	132
University or Master's degree	0.118	101	118
Total	**1**	**1000**	**1000**

… continuation …

Calculations

- $\chi^2 = 14.29$
- $df = 7 - 0 - 1$
- $c_u = \chi^2_{1-0.05}(6) = \text{qchisq}(0.95{,}6) = 12.591$
- p-value = $1 - \text{pchisq}(14.29{,}6) = 0.0266 = 2.66\%$

Decisions

- H_0 is rejected at 5% → sample is not representative
- H_0 is not rejected at 1% → sample can be considered as representative
- basically it's an inconclusive test as the decision is dependent on the significance level

Test of Uniformity

A food delivery company surveyed 100 randomly selected customers to determine their preferred delivery time between 9 AM and 2 PM.

We aim to test whether the distribution of delivery requests follows a uniform distribution at the 5% significance level ($\alpha = 0.05$).

Time Slot	Number of Delivery Requests
09.00 – 10.00	12
10.00 – 11.00	18
11.00 – 12.00	26
12.00 – 13.00	24
13.00 – 14.00	20
Total	**100**

Calculations

- p-value = 19.91%

Decisions

- H_0 is not rejected at any common significance level
- the distribution of delivery requests can be considered uniform
- no need to allocate extra couriers to specific time slot

Time Slot	n_i	n_i^*
09.00 – 10.00	12	20
10.00 – 11.00	18	20
11.00 – 12.00	26	20
12.00 – 13.00	24	20
13.00 – 14.00	20	20
Total	**100**	**100**

Test of Normality

In air transportation, monitoring the average passenger weight is crucial to prevent overloading and optimize capacity utilization. One airline designs its weight load calculations based on an average passenger weight of 80 kg with a standard deviation of 10 kg. To verify this assumption, the weights of randomly selected passengers were measured.

Weight (kg)	Number of Passengers
< 60	13
60 – 70	39
70 – 80	84
80 – 90	70
90 – 100	33
> 100	11

- Testing the mean is possible but for the standard deviation we need to assume that the weights follow a normal distribution
- Null Hypothesis (H_0):
  - assume that the population follows a normal distribution with parameters μ = 80 and σ = 10
- Calculations
  - theoretical probabilities are derived from the cumulative distribution function (CDF)
    - $F(x) = P(weight < x)$
  - the CDF value at the last interval should be 1
  - probabilities are calculated as differences between CDF values
    - e.g. $P(60 < weight < 70) = P(weight < 70) - P(weight < 60) = F(70) - F(60)$
  - expected frequencies are computed as the product of probabilities and sample size

… continuation …

Further Calculations

- the test conditions are met because all expected frequencies exceed 5
- chi-square statistic: $\chi^2 = 17.91$
  - degrees of freedom: $df = 6 - 1$ → value of the test statistics „seems large"
- p-value: 0.31%
- the p-value is quite small, indicating statistical significance

Lower Bound	Upper Bound	n_i	$F(x)$	P_i	n_i^*
	60	13	0.0228	0.0228	5.69
60	70	39	0.1587	0.1359	33.98
70	80	84	0.5000	0.3413	85.34
80	90	70	0.8413	0.3413	85.34
90	100	33	0.9773	0.1359	33.98
100		11	1	0.0228	5.69

… continuation …

Decision

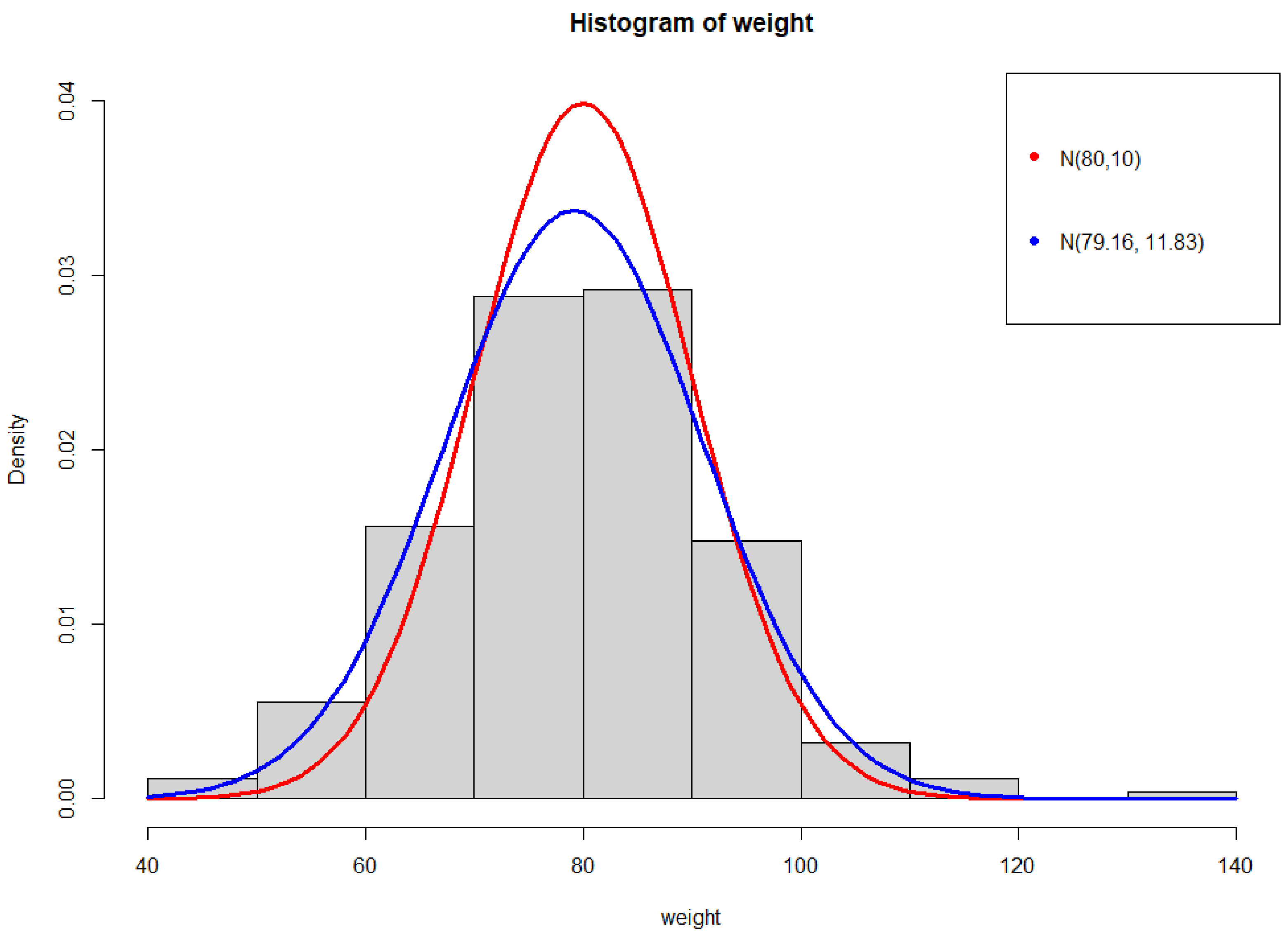
- at common significance levels, H_0 is rejected
- this does not imply the data is not normally distributed, but rather that it does not follow an N(80,10) distribution
- the issue may not be with normality but with the parameters (μ and σ)
- re-estimate μ and σ from the sample:
  - sample mean: $\bar{y} = 79.16$
  - sample standard deviation: $s = 11.83$
  - number of estimated parameters: $b = 2$

… continuation …

- Calculations for N(79.16, 11.83)
  - chi-square statistic: $\chi^2 = 1.18$
    - degrees of freedom: $df = 6 - 2 - 1 = 3 \rightarrow$ statistic appears reasonably small and is on the left tail
  - p-value: 75.77%
    - p-value is greater than 50%, indicating no significant deviation
- Decision: H_0 is not rejected → The data can be considered normally distributed

Lower Bound	Upper Bound	n_i	$F(x)$	P_i	n_i^* N(79.16,11.83)	n_i^* N(80,10)
	60	13	0.0527	0.0527	13.18	5.69
60	70	39	0.2195	0.1667	41.68	33.98
70	80	84	0.5283	0.3088	77.21	85.34
80	90	70	0.8202	0.2919	72.97	85.34
90	100	33	0.9609	0.1407	35.18	33.98
100		11	1	0.0391	9.78	5.69

… continuation …



… continuation …

Test for Mean (μ = 80)

- z-score: $z = -1.12$
- p-value: 26.2%
  - since the z-score is negative, we examine the left tail in a two-tailed test (multiply by 2)
- the p-value is large, meaning μ = 80 is plausible

Test for Standard Deviation (σ = 10)

- chi-square statistic: $\chi^2 = 348.7$
- p-value: 0.006%
- the p-value is extremely small, meaning σ = 10 is rejected

Conclusion

- the data follows a normal distribution, but the assumption of σ = 10 is incorrect
- μ = 80 is a reasonable assumption based on the sample
- the airline should update its standard deviation estimate to better reflect passenger weight variability.

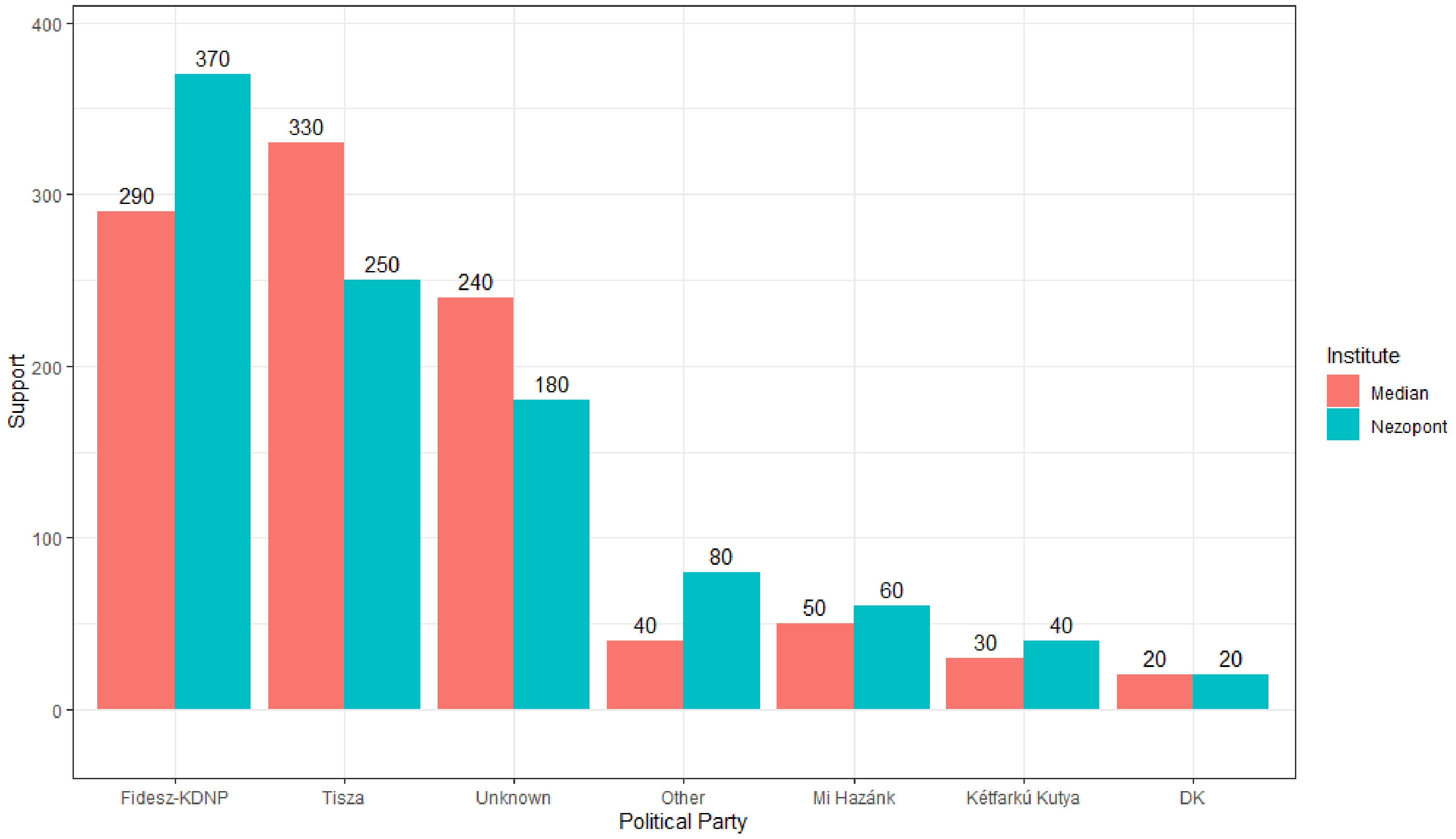
Test of Homogenity

- Testing the Equality of the Distributions of Two Populations
  - the two populations are independent
- Null Hypothesis (H_0)
  - the two distributions are identical
- Frequencies of both samples (X, Y) is analyzed based on the examined variable.

Value	n_{X_i}	n_{Y_i}
C_1	n_{X_1}	n_{Y_1}
C_2	n_{X_2}	n_{Y_2}
…	…	…
C_k	n_{X_k}	n_{Y_k}
Total	n_X	n_Y

- Test Stat: $\chi^2 = n_X n_Y \sum_{i=1}^{k} \frac{1}{n_{X_i}+n_{Y_i}} \cdot \left(\frac{n_{X_i}}{n_X} - \frac{n_{Y_i}}{n_Y}\right)^2$
- Distribution: right-tailed $\chi^2(k-1)$

Can the political opinion polls conducted by Nézőpont and Medián institutes in March 2025 be considered identical based on public opinion polls at a 5% significance level?



Calculations

- $n_{Néző} = n_{Medián} = 1000$
- $\chi^2 = 44.974$
- p-value = 4.7×10^{-8}

Decision

- H_0 is rejected
- they are significantly different