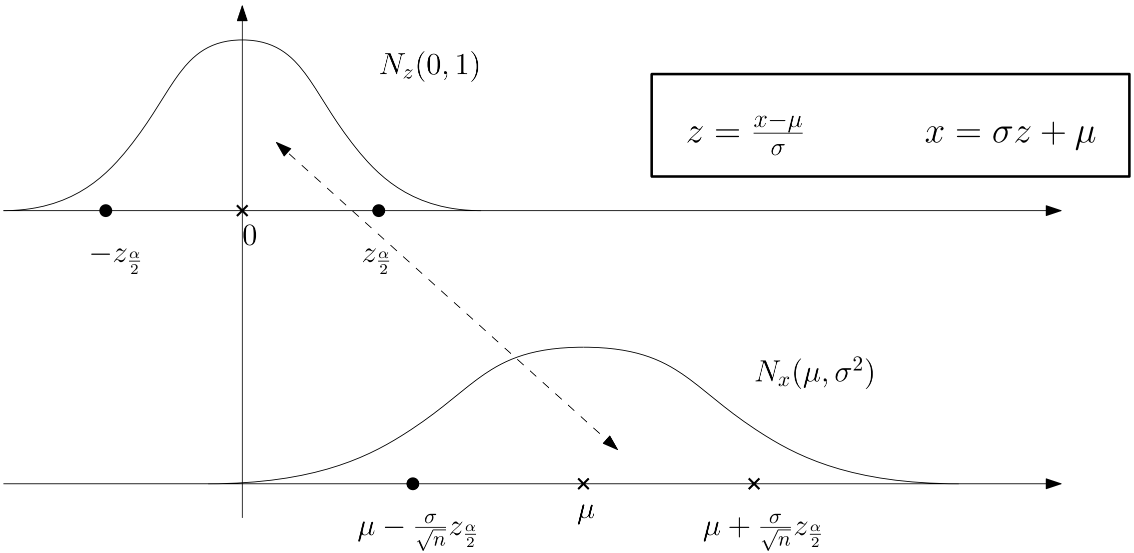
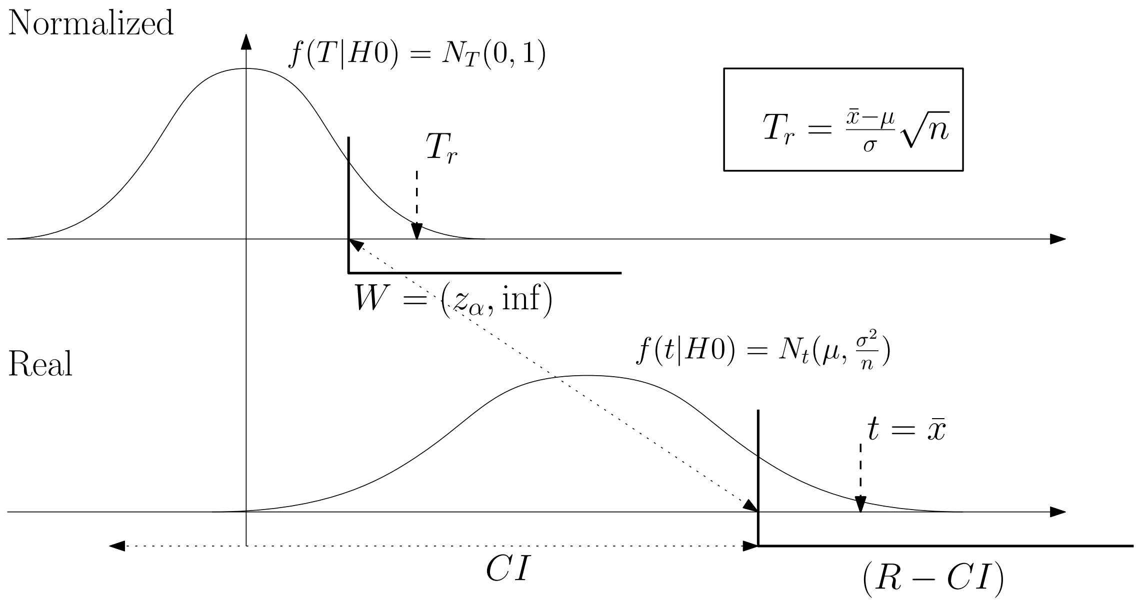
1 Normalization

For $\bar{x} \to \mu$

Normalization



Normalization in test



W is critical region. If $T_r \in W$: reject H0.

2 Overview of tests

How to determine a proper test in Statext

- parametric or nonparametric
- number of variables: 1, 2, more
- tested parameter (property)
- important: H_0 and p-value

Tests with one sample

Parametric tests (normality required)

- `expectation` (known $\times$ unknown variance)
- `proportion`
- `variance`

Nonparametric tests (normality is not required)

- `Wilcoxon test`: tests median

Tests of distribution type

- `w/s test` of normality
- `Kolmogorov-Smirnov test`: tests given distribution.
- `Chi-square test of homogeneity`: test of distribution type.

Tests with two samples

Parametric

- `two expectations`: independent × paired
- `two proportions`
- `two variances`

Nonparametric

- `Mann-Whitney test`: equality of two medians (independent samples)
- `Wilcoxon test`: two medians (paired samples)
- `McNemar test`: improvement after some action. Binary data.

Tests with more samples

Parametric

- `Analysis of variance`: equality of several expectations
- `Anova with two factors`: equality in columns and rows
- `Bartlett` - equality of more variances
- `Scheffé` - detects different expectations

Nonparametric

- `Kruskal-Wallis`: nonparametric anova.
- `Friedman - block test` of equality of medians

Tests of independence

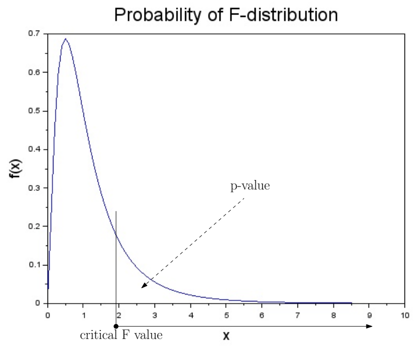
- `Lambda coefficient`: association of two discrete rvs.
- `Pearson test`: independence of two rvs (parametric).
- `Spearman test`: nonparametric Pearson.
- `Chi-square independence`: test of independence of two rvs.

3 Description of tests

F-test

Compares explained variance V_E and unexplained variance V_U. The statistics is

$$F = \frac{V_E}{V_U} \sim F \text{ Fisher distribution}$$



If $F = 0$, p-value=1 - $V_E = 0$ nothing is explained.

If $F \to \infty$, p-value $\to 0$ - $V_U \to 0$ all is explained.

ANOVA I

We have data from several sources (populations). We test, if the expectations of the populations are equal.

The data are $x$es in the following table.

X_1	X_2	X_3
x	x	x
x	x	x
x	x	x
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$
v_1	v_2	v_3

We compute averages $\bar{x}$ and variances v. Then:

- Average of variances v_i corresponds to **unexplained variance** V_U - it describes the overall variance in the data.
- Variance of the averages $\bar{x}_i$ corresponds to **explained variance** V_E - it expresses the variance between classes.

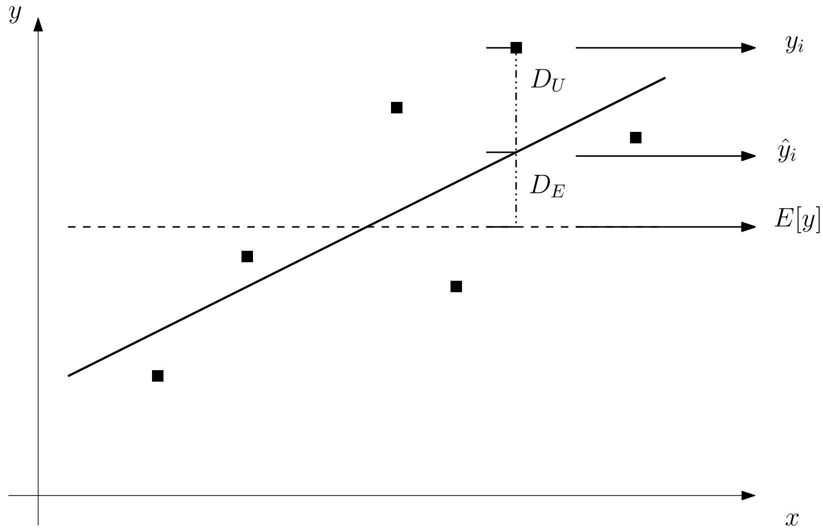
If the explained variance V_E is sufficiently larger with respect to the unexplained one V_U then we conclude that the classes are not equal.

Statistics: $F = \frac{V_E}{V_U} \sim F$ distribution (right-sided test)

ANOVA II

	X_1	X_2	X_3		
Y_1	x	x	x	$\bar{y}_1$	v_{y1}
Y_2	x	x	x	$\bar{y}_2$	v_{y2}
Y_3	x	x	x	$\bar{y}_3$	v_{y3}
Y_4	x	x	x	$\bar{y}_4$	v_{y4}
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$		
	v_{x1}	v_{x2}	v_{x3}		

Regression



Friedman test

Example

A quality of five shops is checked by three evaluators. Each evaluator rates each shop. The results are in the table

Evaluator/shop	1	2	3	4	5
1	x	x	x	x	x
2	x	x	x	x	x
3	x	x	x	x	x

We are interested in evaluation of the shops (not evaluators).

Remark: The shops are called treatment, the evaluators are subjects. In the Statext, we choose "Each data set is for subject" which means, the data in rows come from individual subjects.

Chi-square test

Works for discrete data or continuous ones discretized on intervals.

I based on

- O observed absolute frequencies - from measured data,
- E expected absolute frequencies - constructed so that:
  - H_0 is precisely fulfilled,
  - number of data (sum of frequencies) is the same as for measured ones.

Statistics with χ^2 distribution is

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

Example

Test if the number of traffic accidents is uniformly distributed during the week, if we measured the following number of accidents

Weekdays	Saturday	Sunday
587	98	103
O_1	O_2	O_3

Sum $\sum O_i = 788$. Time axis has 7 intervals (days). Weekdays has 5, Sat and Sun have 1. So we need to divide 788 into groups with proportion 5/1/1.

$$E_1 = \tfrac{788}{7}5 = 562.86, \quad E_2 = \tfrac{788}{7}1 = 112.57, \quad E_3 = \tfrac{788}{7}1 = 112.57$$

$$\chi^2 = 3.73; \quad pv = 0.155$$

As $pv > 0.05$ the H_0 is not rejected at the confidence level $\alpha = 0.05$.

Pearson and Spearman tests

Pearson test tests the correlation coefficient

$$R = \frac{C\left[X,Y\right]}{\sqrt{D\left[X\right]D\left[Y\right]}}.$$

If $R = 0$ the random variables X and Y are uncorrelated. If $R \to -1$ or 1, the variables are strongly correlated. The test is both sided.

Its statistics is the sample correlation coefficient

$$r = \frac{cov\left(x,y\right)}{\sqrt{var\left(x\right)var\left(y\right)}} \sim t \text{ Student distribution}$$

H0: The variables are independent.

For p-value small, the independence is rejected - variables are dependent.

Spearman test is a nonparametric variant of Pearson. Instead of data it uses their ranks. Can be used for discrete variables.

McNemar test

We measure binary (yes/not) data before and after some action. We test if the action caused any change.

Example

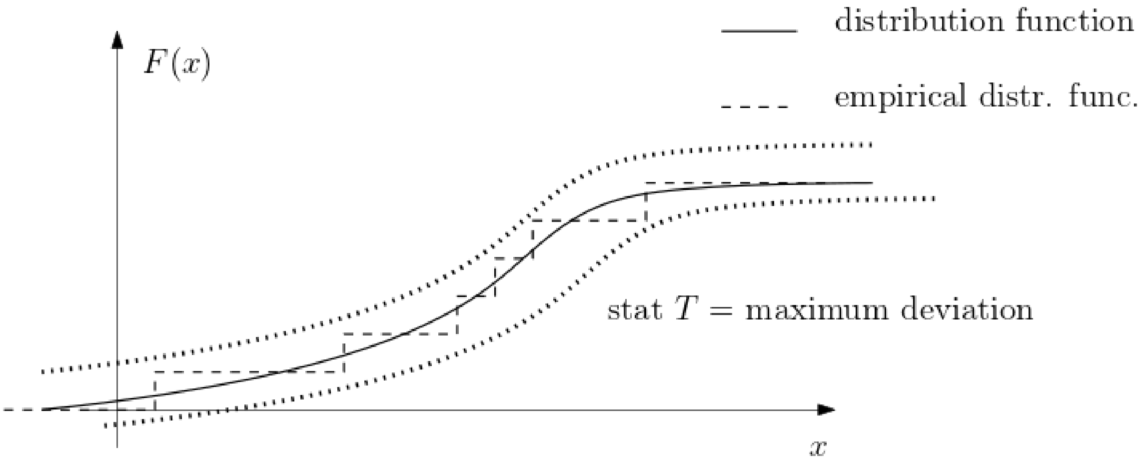
We ask 20 people if they have a cold. The answers are yes/not. Then we give them some medicine and ask again. The question is whether there was any change after the application of the drug (either positive or negative).

H0: no change.

$\rightarrow$ if p-value is small, a change has been detected.

Kolmogorov Smirnov test

Tests type of distribution.



4 Validation in regression

- Graph xy
- Test of independence x and y (Pearson, Spearman).
- F test of prediction.
- Independence and autocorrelation of residuals $(e_t = ae_{t-1} + \epsilon_t)$
- Prediction error $RPE = \text{var}\,(y - y_p)\,/\text{var}\,(y)$.