HELP TO STATISTICAL FUNCTIONS

Contents

1 Intervals of confidence

– CI for expectation (known variance)

[lb, ub]=z_int (av,vr,n,side,alpha)

– CI for expectation (unknown variance)

[lb, ub]=t_int (av,vr,n,side,alpha)

– interval for proportion

[lb, ub]=prop_int(p,n,side,alpha)

– interval for variance

[lb, ub]=var_int(vr,n,side,alpha)

2 One sample tests

2.1 Parametric

– test of expectation (known variance)

pv=z_test(Mu0,av,vr,n,side)

– test of expectation (unknown var)

pv=t_test(mu0,av,vr,n,side)

mu0 ··· expectation according to H0

– test of ratio

pv=prop_test(p0,p,n,side)

p0 ··· proportion according to H0

– test of variance

pv=var_test(sig0,vr,n,side)

sig0 ··· variance according to H0

2.2 Nonparametric

– Wilcoxon test of median

pv=wilcoxon_test(x,med0,side)

med0 ··· median according to H0

3 Two sample tests

3.1 Parametric

– test of two expectations (indep. samples, equal variances)

pv=t_test_2s(av1,vr1,n1,av2,vr2, n2,side)

– test of two expectations (indep. samples, different variances)

pv=t_test_2n(av1,vr1,n1,av2,vr2, n2,side)

– test of two expectations (paired samples)

pv=t_test_2p(x1,x2,side)

– test of two ratios

pv=prop_test_2(p1,n1,p2,n2,side)

– test of two variances

pv=var_test_2(vr1,n1,vr2,n2,side)

3.2 Nonparametric

– Mann Whitney test of two medians (indep. samples)

pv=mannwhit_test(x1,x2)

– Wilcoxon paired test of two medians (paired samples)

pv=wilcoxon_test(x1,x2,side)

– McNemar test of a change after action (dichotomous samples)

pv=mcnemar_test(KT)

KT ··· frequency table 2×2

4 More sample tests

4.1 Parametric

– ANOVA (one-way)

pv=anova_1(T)

T ··· matrix with samples in columns

– ANOVA (two-way)

[pv_col, pv_row]=anova_2(T)

T ··· matrix with data

– Bartlett test of several variances

pv=bartlett_test(T)

T ··· matrix with samples in columns

4.2 Nonparametric

– Kruskal-Wallis test of several medians (indep. samples)

pv=kruskal_test(T)

T ··· matrix with samples in columns

– Friedman test of several medians (paired, block test)

pv=friedman_test(T)

T ··· matrix with samples in columns

5 Tests of distribution type

– chi square test of homogeneity

pv=chisquare_test(O,E)

– chi2 test of normality

pv=normCh2_test(x,n)

n ··· number of intervals

– Shapiro test of normality

pv=shapiro_test(x)

– Kolmogorov - Smirnov test for continuous data

pv=ks_test_cont(F)

F ··· vector of values of assumed distribution function in measured values

– Kolmogorov - Smirnov test for discrete data

pv=ks_test_disc(F,Fm1)

F ··· vector of values of assumed distribution function in measured values

Fm1 ··· vector of values of assumed distribution function in measured values reduced by one. The domain of natural or whole numbers is assumed.

– Kolmogorov - Smirnov test for two samples

pv=ks_test_2(x1,x2)

6 Tests of independence

– Pearson test of independence

pv=pearson_test(x1,x2)

– Spearman test of independence

pv=spearman_test(x1,x2)

– chi2 test of independence

pv=chisquare_test_i(KT)

KT ··· contingency (frequency) table

7 Regression analysis

7.1 Linear

– linear regression

par=lin_reg(x,y)

– linear prediction

yp=lin_pred(x,par)

– multivariate linear regression

par=lin_reg_n(x,y)

– multivariate linear prediction

yp=lin_pred_n(x,par)

7.2 Nonlinear

– exponential regression

par=exp_reg(x,y)

– exponential prediction

yp=exp_pred(x,par)

– polynomial regression

par=pol_reg(x,y,k)

k ··· order of the polynomial

– polynomial prediction

yp=pol_pred(x,par)

7.3 Validation

– characteristics of linear regression

[b1,b0,r]=reg_desc(x,y)

– F test of prediction

pv=f_test_pred(y,yp,np)

np ··· number of parameters in regression

– test of independence of sequence elements (residuals)

pv=ordinal_test(x)

– test of autocorrelation in a sample

res=autoreg_test(x,y)

8 Auxiliary functions

– close all figures

cla

– combination number n!/(k'*(n-k)!) computed in logarithms

c=comb(n,k)

– covariance of x,y (second moment)

c=cvar(x,y)

– logarithm of factorial

Lf=factL(n)

– conversion of list to matrix

M=list2mat(L)

– list from columns of matrix

L=mat2list(M)

– ranks of a discrete random variable

i.e. for a=[3 5 2 2] we get j=[3 4 1.5 1.5]

j=ranks(a)

– sampling without repetition

n ··· *sample length*

a ··· *set to be sampled*

s=samp(n,a)

– scatter plot

scatt(x,y)

– frequency table

T = table(x, y)

– different values of a variable and their frequencies

v=vals(x)

v ··· [values; abs.freq.]

– variance (second central moment)

v=var(x)

– average computed from values and frequencies

m=meanf(x,f)

x ··· *values, f* ··· *frequencies*

– variance computed from values and frequencies

v=varf(x,f)

x ··· *values, f* ··· *frequencies*

9 Common notation

x, x1, x2 ··· sample

y ··· dependent sample

av ··· sample average

vr ··· sample variance

side ··· test (interval) orientation
'l' - left, 'r' - right, 'b' - both

alpha ··· confidence level

p ··· proportion

n ··· sample length

T $\cdots$ matrix of samples in columns
(for samples with different lengths use list)

O $\cdots$ observed absolute frequencies

E $\cdots$ expected absolute frequencies

par $\cdots$ parameters (of regression)