

**JSP: A J Statistical Package
Second Edition**

**Keith Smillie
Department of Computing Science
University of Alberta
Edmonton, Alberta T6G 2H1
smillie@cs.ualberta.ca**

Introduction

JSP is a package of statistical programs written in **J4.01** with the main statistical verbs in the following categories: Summarization; Frequencies; Correlation and regression; Analysis of variance; Chi-square; Nonparametric statistics; Simulation; Probability distributions. The script file is available by anonymous ftp at *ftp.cs.ualberta.ca* in the file *pub/smillie/jsp.ijs*.

The documentation of almost all of the verbs has the following format:

name Left argument, if any (Integers **m**, **n**; **u**, **v** integer or real; lists **x**, **y**;
Right argument table **t**)
Explicit result

Summarization

am -
y
Arithmetic mean of **y**

gm -
y
Geometric mean of **y**

var -
y
Variance of **y**

sd -
y
Standard deviation of **y**

Q1 -
y
First quartile of **y**

Q2 or median -
y
Median of **y**

Q3 -
y
Third quartile of **y**

five -
y
Min., first quartile, median, third
quartile and max. of **y**

summary -
y
Summary statistics (with labels) of **y**

mode -
y
Mode of **y**

Frequencies

fr **x** Range
y (Integer)
Frequencies over range of **y**

frtab **x** Range
y (Integer)
Frequency table over range of **y**

nubfr -
y (Integer)
Nub frequencies of **y**

nubtab -
y (Integer)
Frequency table over nub of **y**

bnumtab -
y
Boxed frequency table over nub of **y**

cfr **x** (End points of classes)
y (Integer or real)
Class frequencies

cfrtab **x** (As in **cfr**)
y (Integer)
Frequency table with mid-points in 1st col. and frequencies in 2nd

FR **x** (List of pairs giving ranges for each axis)
y (Integer pairs)
Two-way freq. table over range of **y**

barchart **x** (Range)
y (Frequencies)
Range in 1st col. and frequencies as * in 2nd.

vbarchart **x** (Range)
y (Frequencies)
Frequencies as * given vertically with no range

SLtab -
y (Integer)
Frequency table of stems

stemtab -
y (Integer)
Stem-and-leaf table

Correlation and regression

cov **x**
y
Covariance between **x** and **y**

cor **x**
y
Correlation coefficient between **x** and **y**

covtab -
List, each item of which is a list
Variance-covariance table of all pairs

cortab -
List, each item of which is a list
Correlation table of all pairs

SR **x**
y
Simple linear regression with dep. var. **y** and indep. var. **x**

REG -
List, each item of which is a list with last item dep. variable
Multiple linear regression

Analysis of variance

AOV [**x**]
Table or higher-dimensional array
AOV table with **x** giving specified terms (Default gives all terms, e.g., **AOV t** is '**A B AB**' **AOV t**)

aov1 -
List, each item of which gives observations for one level in one-way AOV with unequal subclass numbers
AOV table

Chi-square

ExpFrTab -
t (Obs. freq.)
t (Exp. freq.)

chisq **x** or **t** (Obs. freq.)
x or **t** (Exp. freq.)
Chi-square

chisq22 -
t (Obs. freq. for 2-by-2 table)
Probability

Nonparametric statistics

uranks -
y
Ranks of items of **y** unadjusted for ties

ranks -
y
Ranks with ties averaged

invranks -
y

Ranks in inverse order

rcor **x**
y
Rank correlation coefficient between **x**
and **y**

runs -
y
Number of runs

Simulation

Die -
n
Results of rolling die **n** times

Dice **m**
n
Results of rolling **m** dice **n** times as an
n-item list with **m** items in each item

SumDice **m**
n
n-item list of sums

Heads **m**
n
n-item list of no. of heads

rand -
n, x, ...
Uniformly distributed random numbers,
e.g., **rand 3** is a 3-item list,
rand 2 4 is 2 by 4 table

nmlrand [**u,v**]
n
n normal deviates with mean **u** and s.d.
v. Default is standard normal

exprand **m**
n
Exponentially distributed random
numbers with mean **m**

Probability distributions

binomial **n, p** (Number of trials and prob. of
success in a single trial)
m or **y** (Number of successes)
Probabilities

poisson **m** (Mean)
n or **y** (Number of successes)

Probabilities

geometric **p** (Probability of success in a single
trial)
n or **y** (Number of trials)
Probabilities

hgeometric **x** (3-item list giving no. in population
of type A, no. of type not-A, sample
size)
n, x (No. in sample of type A)
Hypergeometric probabilities

ndistn -
u or **y**
Prob. density function values

tdistn **m** (Degrees of freedom)
u or **y**
Prob. density function values

chisq **m** (Degrees of freedom)
u or **y**
Prob. density function values

fdistn **m,n** (Num. & denom. d.f.)
u or **y**
Prob. density function values

Cumulative probabilities for the last four
distributions may be found by means of the integral
adverb **I**, Iverson (1993), and typical uses are as
follows:

```
ndistn I 0 1 2 3
5&tdistn I 2.015 2.571 3.365
10&csdistn I 12.5 16 18.3
5 20&fdistn I 2.16 2.71 3.29 4.1
```

References

Iverson, K. E., 1993. *Calculus*. Iverson Software Inc.,
Toronto.

Smillie, Keith, 1999. *J Companion for Statistical
Calculations*.

