196

DBMS_STAT_FUNCS

The `DBMS_STAT_FUNCS` package provides statistical functions.

This chapter contains the following topic:

- Summary of DBMS_STAT_FUNCS Subprograms

Summary of DBMS_STAT_FUNCS Subprograms

This table lists and briefly describes the DBMS_STAT_FUNCS subprograms.

Table 196-1 DBMS_STAT_FUNCS Package Subprograms

Subprogram	Description
EXPONENTIAL_DIST_FIT Procedure	Tests how well a sample of values fits an exponential distribution
NORMAL_DIST_FIT Procedure	Tests how well a sample of values fits a normal distribution
POISSON_DIST_FIT Procedure	Tests how well a sample of values fits a Poisson distribution
SUMMARY Procedure	Summarizes a numerical column of a table
UNIFORM_DIST_FIT Procedure	Tests how well a sample of values fits a uniform distribution
WEIBULL_DIST_FIT Procedure	Tests how well a sample of values fits a Weibull distribution

EXPONENTIAL_DIST_FIT Procedure

This procedure tests how well a sample of values fits an exponential distribution.

Syntax

```
DBMS_STAT_FUNCS.EXPONENTIAL_DIST_FIT (
   ownername    IN    VARCHAR2,
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   test_type    IN    VARCHAR2 DEFAULT 'KOLMOGOROV_SMIRNOV',
   lambda       IN    NUMBER,
   mu           IN    NUMBER,
   sig          OUT   NUMBER);
```

Parameters

Table 196-2 *EXPONENTIAL_DIST_FIT Procedure Parameters*

Parameter	Description
`ownername`	The schema where the table resides.
`tablename`	The table where the column resides.
`columnname`	The column of the table against which to run the test.

Table 196-2 (Cont.) *EXPONENTIAL_DIST_FIT Procedure Parameters*

Parameter	Description
test_type	The type of test to use: 'CHI_SQUARED', 'KOLMOGOROV_SMIRNOV' or 'ANDERSON_DARLING'.
lambda	The scale parameter.
mu	The location parameter.
sig	The goodness of fit value, based on test type. A small value indicates a significant difference between the sample and the exponential distribution. A number close to 1 indicates a close match.

NORMAL_DIST_FIT Procedure

This procedure tests how well a sample of values fits a normal distribution.

Syntax

```
DBMS_STAT_FUNCS.NORMAL_DIST_FIT (
   ownername    IN    VARCHAR2,
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   test_type    IN    VARCHAR2 DEFAULT 'SHAPIRO_WILKS',
   mean         IN    NUMBER,
   stdev        IN    NUMBER,
   sig          OUT   NUMBER);
```

Parameters

Table 196-3 NORMAL_DIST_FIT Procedure Parameters

Parameter	Description
ownername	The schema where the table resides.
tablename	The table where the column resides.
columnname	The column of the table against which to run the test.
test_type	The type of test to use: 'CHI_SQUARED', 'KOLMOGOROV_SMIRNOV', 'ANDERSON_DARLING' or 'SHAPIRO_WILKS'.
mean	The mean of the distribution against which to compare.
stdev	The standard deviation of the distribution against which to compare.
sig	The goodness of fit value, based on test type. A small value indicates a significant difference between the sample and the normal distribution. A number close to 1 indicates a close match.

POISSON_DIST_FIT Procedure

This procedure tests how well a sample of values fits a Poisson distribution.

Syntax

```
DBMS_STAT_FUNCS.POISSON_DIST_FIT (
   ownername    IN    VARCHAR2,
```

```
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   test_type    IN    VARCHAR2 DEFAULT 'KOLMOGOROV_SMIRNOV',
   lambda       IN    NUMBER,
   sig          OUT   NUMBER);
```

Parameters

Table 196-4 POISSON_DIST_FIT Procedure Parameters

Parameter	Description
ownername	The schema where the table resides.
tablename	The table where the column resides.
columnname	The column of the table against which to run the test.
test_type	The type of test to use: 'KOLMOGOROV_SMIRNOV' or 'ANDERSON_DARLING'.
lambda	The lambda parameter is the shape parameter.
sig	The goodness of fit value, based on test type. A small value indicates a significant difference between the sample and the Poisson distribution. A number close to 1 indicates a close match.

SUMMARY Procedure

This procedure summarizes the numerical column specified in the `columnname` of `tablename`.

The summary is returned as a Summary Type. Note that most of the output of `SUMMARY` can be obtained with currently available SQL.

Syntax

```
DBMS_STAT_FUNCS.SUMMARY (
   ownername    IN    VARCHAR2,
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   sigma_value  IN    NUMBER DEFAULT 3,
   s           OUT    SummaryType);
```

Parameters

Table 196-5 SUMMARY Procedure Parameters

Parameter	Description
ownername	The schema where the table resides.
tablename	The table where the column resides.
columnname	The column of the table to be summarized.
sigma_value	The number of sigmas for the set of extreme values, defaults to 3.
s	The Record containing summary information about given column.

Definition of SummaryType

```
TYPE n_arr IS VARRAY(5) of NUMBER;
TYPE num_table IS TABLE of NUMBER;
```

```
TYPE summaryType IS RECORD (
   count             NUMBER,
   min               NUMBER,
   max               NUMBER,
   range             NUMBER,
   mean              NUMBER,
   cmode             num_table,
   variance          NUMBER,
   stddev            NUMBER,
   quantile_5        NUMBER,
   quantile_25       NUMBER,
   median            NUMBER,
   quantile_75       NUMBER,
   quantile_95       NUMBER,
   plus_x_sigma      NUMBER,
   minus_x_sigma     NUMBER,
   extreme_values    num_table,
   top_5_values      n_arr,
   bottom_5_values   n_arr);
```

UNIFORM_DIST_FIT Procedure

This procedure tests well a sample of values fits a uniform distribution.

Syntax

```
DBMS_STAT_FUNCS.UNIFORM_DIST_FIT (
   ownername    IN    VARCHAR2,
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   var_type     IN    VARCHAR2 DEFAULT 'CONTINUOUS',
   test_type    IN    VARCHAR2 DEFAULT 'KOLMOGOROV_SMIRNOV',
   paramA       IN    NUMBER,
   paramB       IN    NUMBER,
   sig          OUT   NUMBER);
```

Parameters

Table 196-6 UNIFORM_DIST_FIT Procedure Parameters

Parameter	Description
ownername	The schema where the table resides.
tablename	The table where the column resides.
columnname	The column of the table against which to run the test.
var_type	The type of distribution: 'CONTINUOUS' (the default) or 'DISCRETE'
test_type	The type of test to use: 'CHI_SQUARED', 'KOLMOGOROV_SMIRNOV' or 'ANDERSON_DARLING'.
paramA	Parameter A estimated from the sample (the location parameter).
paramB	Parameter B estimated from the sample (the scale parameter).
sig	The goodness of fit value, based on test type. A small value indicates a significant difference between the sample and the uniform distribution. A number close to 1 indicates a close match.

WEIBULL_DIST_FIT Procedure

This procedure tests how well a sample of values fits a Weibull distribution.

Syntax

```
DBMS_STAT_FUNCS.WEIBULL_DIST_FIT (
   ownername    IN    VARCHAR2,
   tablename    IN    VARCHAR2,
   columnname   IN    VARCHAR2,
   test_type    IN    VARCHAR2 DEFAULT 'KOLMOGOROV_SMIRNOV',
   alpha        IN    NUMBER,
   mu           IN    NUMBER,
   beta         IN    NUMBER,
   sig          OUT   NUMBER);
```

Parameters

Table 196-7 WEIBULL_DIST_FIT Procedure Parameters

Parameter	Description
ownername	The schema where the table resides.
tablename	The table where the column resides.
columnname	The column of the table against which to run the test.
test_type	The type of test to use: 'CHI_SQUARED', 'KOLMOGOROV_SMIRNOV' or 'ANDERSON_DARLING'.
alpha	The scale parameter.
mu	The location parameter.
beta	The slope/shape parameter.
sig	The goodness of fit value, based on test type. A small value indicates a significant difference between the sample and the Weibull distribution. A number close to 1 indicates a close match.