Балтийский государственный технический университет «Военмех» им. Д.Ф.Устинова

Кафедра прикладной математики и информатики (И7).

Математическая статистика.

Лабораторная работа №6.

Критерии согласия в статистическом пакете Statgraphics.

Вариант 12.

Выполнил

Романов Н.

группа И383

Проверил

Гнидин В.В.

Санкт-Петербург

2011

**Задание 1.**

Смоделировать выборку объемом 50 элементов согласно гипотетическому логнормальному распределению с параметрами m = 8 и σ = 12. Проверить с помощью критериев согласия соответствие статистического и гипотетического распределений с уровнем значимости α = 0.1.

**Решение в пакете Statgraphics:**

Смоделированная выборка:

**Probability Distributions**

Distribution: Lognormal

Parameters: Mean Std. dev.

Dist. 1 8 12

**The StatAdvisor**

*This procedure allows you to analyze any of 24 probability distributions. Currently, the lognormal distribution has been selected. You can create various plots, compute tail areas and critical values, and generate random numbers from the selected distribution. Up to five sets of parameters can be specified by pressing the alternate mouse button and selecting Analysis Options.*

**Cumulative Distribution**

Distribution: Lognormal

Lower Tail Area (<)

Variable Dist. 1

0 0,0

Probability Density

Variable Dist. 1

0 0,0

Upper Tail Area (>)

Variable Dist. 1

0 1,0

**The StatAdvisor**

*This pane evaluates the cumulative lognormal distribution. It will calculate the tail areas for up to 5 critical values of the distribution. It will also calculate the probability density or mass function. For example, the output indicates that, for the first distribution specified, the probability of obtaining a value less than 0,0 is 0,0. Also, the probability of obtaining a value greater than 0,0 is 1,0. The height of the probability density function at 0,0 is 0,0.*

**Inverse CDF**

Distribution: Lognormal

CDF Dist. 1

0,01 0,355045

0,1 1,10383

0,5 4,4376

0,9 17,84

0,99 55,4643

**The StatAdvisor**

*This pane finds critical values for the lognormal distribution. You may specify up to 5 five tail areas. The critical value is defined as the largest value for the lognormal distribution such that the probability of not exceeding that value does not exceed the area specified. For example, the output indicates that, for the first distribution specified, 0,355045 is the largest value such that the probability of not exceeding 0,355045 is less than or equal to 0,01.*

**Random Numbers**

To generate random numbers from the selected distribution,

use the save button on the analysis toolbar.

Random numbers to be generated: 50

**The StatAdvisor**

* This pane allows you to specify the number of observations desired in a random sample from the lognormal distribution. You set the number of observations by pressing the alternate mouse button and selecting Pane Options. After setting the size, press the Save Results button on the analysis toolbar. This allows you to save random samples from the specified distribution in columns of the current data file. Every time you select Save Results, a new random sample will be generated.*



Проверка соответствия:

**Analysis Summary**

Data variable: lognormal

100 values ranging from 0,36691 to 85,1661

Fitted normal distribution:

mean = 9,08297

standard deviation = 13,8584

**The StatAdvisor**

*This analysis shows the results of fitting a normal distribution to the data on lognormal. The estimated parameters of the fitted distribution are shown above. You can test whether the normal distribution fits the data adequately by selecting Goodness-of-Fit Tests from the list of Tabular Options. You can also assess visually how well the normal distribution fits by selecting Frequency Histogram from the list of Graphical Options. Other options within the procedure allow you to compute and display tail areas and critical values for the distribution. To select a different distribution, press the alternate mouse button and select Analysis Options.*

**Tests for Normality for lognormal**

Computed Chi-Square goodness-of-fit statistic = 229,28

P-Value = 0,0

Shapiro-Wilks W statistic = 0,604174

P-Value = 0,0

Z score for skewness = 5,38448

P-Value = 7,2817E-8

Z score for kurtosis = 5,6361

P-Value = 1,74419E-8

**The StatAdvisor**

*This pane shows the results of several tests run to determine whether lognormal can be adequately modeled by a normal distribution. The chi-square test divides the range of lognormal into 24 equally probable classes and compares the number of observations in each class to the number expected. The Shapiro-Wilks test is based upon comparing the quantiles of the fitted normal distribution to the quantiles of the data. The standardized skewness test looks for lack of symmetry in the data. The standardized kurtosis test looks for distributional shape which is either flatter or more peaked than the normal distribution.*

*The lowest P-value amongst the tests performed equals 0,0. Because the P-value for this test is less than 0.01, we can reject the idea that lognormal comes from a normal distribution with 99% confidence.*

**Goodness-of-Fit Tests for lognormal**

Chi-Square Test

----------------------------------------------------------------------------

Lower Upper Observed Expected

Limit Limit Frequency Frequency Chi-Square

----------------------------------------------------------------------------

at or below -6,85908 0 12,50 12,50

-6,85908 -0,264418 0 12,50 12,50

-0,264418 4,66711 58 12,50 165,62

4,66711 9,08297 18 12,50 2,42

9,08297 13,4988 8 12,50 1,62

13,4988 18,4304 3 12,50 7,22

18,4304 25,025 4 12,50 5,78

above 25,025 9 12,50 0,98

----------------------------------------------------------------------------

Chi-Square = 208,641 with 5 d.f. P-Value = 0,0

Estimated Kolmogorov statistic DPLUS = 0,272684

Estimated Kolmogorov statistic DMINUS = 0,264694

Estimated overall statistic DN = 0,272684

Approximate P-Value = 6,95793E-7

EDF Statistic Value Modified Form P-Value

---------------------------------------------------------------------

Kolmogorov-Smirnov D 0,272684 2,7473 <0.01\*

Anderson-Darling A^2 14,2087 14,3185 0,0000\*

---------------------------------------------------------------------

*\*Indicates that the P-Value has been compared to tables of critical values specially constructed for fitting the currently selected distribution. Other P-values are based on general tables and may be very conservative.*

**Tail Areas for lognormal**

area below 7,26637 = 0,447853

area below 8,17467 = 0,473869

area below 9,08297 = 0,5

area below 9,99127 = 0,526131

area below 10,8996 = 0,552147

**The StatAdvisor**

*This pane calculates tail areas for the fitted normal distribution. It will calculate the tail areas for up to 5 critical values, which you may specify by pressing the alternate mouse button and selecting Pane Options. For example, the output indicates that the probability of obtaining a value less than or equal to 7,26637 is 0,447853.*

**Critical Values for lognormal**

area below -23,1566 = 0,01

area below -8,67734 = 0,1

area below 9,08297 = 0,5

area below 26,8433 = 0,9

area below 41,3225 = 0,99

**The StatAdvisor**

*This pane calculates critical values for the fitted normal distribution. It will calculate the critical values for up to 5 lower tail areas, which you may specify by pressing the alternate mouse button and selecting Pane Options. For example, the output indicates that the value of the fitted normal distribution below which you would find an area equal to 0,01 is -23,1566.*

**Normal Tolerance Limits for lognormal**

Normal distribution

Sample size = 100

Mean = 9,08297

Sigma = 13,8584

95,0% tolerance interval for 99,73% of the population

Xbar +/- 3,41743 sigma

Upper: 56,4432

Lower: -38,2772

**The StatAdvisor**

*Assuming that lognormal comes from a normal distribution, the tolerance limits state that we can be 95,0% confidence that 99,73% of the distribution lies between -38,2772 and 56,4432. This interval is computed by taking the mean of the data +/-3,41743 times the standard deviation. This interval is only reliable if the data comes from a normal distribution, which you can test by selecting Tests for Normality from the list of Tabular Options. If the data do not come from a normal distribution, select Distribution-Free Limits from the list of Tabular Options.*

**Distribution-Free Tolerance Limits for lognormal**

Data summary

Count = 100

Maximum = 85,1661

Median = 3,9192

Minimum = 0,36691

95,0% tolerance interval for 95,3433% of the population

Upper: 85,1661

Lower: 0,36691

(Based on an interval depth = 1)

**The StatAdvisor**

*Without assuming that lognormal comes from a normal distribution, the tolerance limits state that we can be 95,0% confidence that 95,3433% of the distribution lies between 0,36691 and 85,1661. This interval is computed from the smallest and largest data values.*



