**Балтийский государственный технический университет**

**«Военмех» им. Д.Ф. Устинова**

**Кафедра И7**

**«Кафедра математической статистики и прикладной математики»**

**«Математическая статистика»**

Лабораторная работа № 5

«Проверка статистических гипотез о числовых значениях нормальных распределений в пакетах STATGRAPHIC и MathCAD»

Вариант 18

**Выполнил:**

Студент Тихонов А.Ю.

Группа И383

**Преподаватель:**

Гнидин В.В.

Санкт-Петербург

2011

**Задание 1.**

Смоделировать нормальную выборку с параметрами Mx = 9 (месяц рождения) и Dx = 18 (номер фамилии по списку) объемом 50 элементов. Смоделировать равномерную выборку и засорить ей нормальную. Смоделировать вторую нормальную выборку с параметрами My = Mx + 1.5 = 10.5, Dy = Dx + 3 = 21 и объемом 100 элементов. Проверить гипотезы о числовых значениях мат.ожидания и дисперсии.

**Решение в пакете Statgraphics**

Нормальная выборка:

**Probability Distributions**

Distribution: Normal

|  |  |  |
| --- | --- | --- |
| *Parameters:* | *Mean* | *Std. Dev.* |
| Dist. 1 | 9 | 4,24 |
| Dist. 2 |  |  |
| Dist. 3 |  |  |
| Dist. 4 |  |  |
| Dist. 5 |  |  |

**The StatAdvisor**

*This procedure allows you to analyze any of 45 probability distributions. Currently, the Normal 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: Normal

Lower Tail Area (<)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 0,0168918 |  |  |  |  |

Probability Density

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 0,00988927 |  |  |  |  |

Upper Tail Area (>)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 0,983108 |  |  |  |  |

**The StatAdvisor**

*This pane evaluates the cumulative Normal. 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 0,0168918. Also, the probability of obtaining a value greater than 0,0 is 0,983108. The height of the probability density function at 0,0 is 0,00988927.*

**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 Normal. 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.*

****

****

****

Равномерная выборка:

**Probability Distributions**

Distribution: Uniform

|  |  |  |
| --- | --- | --- |
| *Parameters:* | *Lower Limit* | *Upper Limit* |
| Dist. 1 | 0 | 0,5 |
| Dist. 2 |  |  |
| Dist. 3 |  |  |
| Dist. 4 |  |  |
| Dist. 5 |  |  |

**The StatAdvisor**

*This procedure allows you to analyze any of 45 probability distributions. Currently, the Uniform 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: Uniform

Lower Tail Area (<)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 0,0 |  |  |  |  |

Probability Density

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 2,0 |  |  |  |  |

Upper Tail Area (>)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Variable* | *Dist. 1* | *Dist. 2* | *Dist. 3* | *Dist. 4* | *Dist. 5* |
| 0 | 1,0 |  |  |  |  |

**The StatAdvisor**

*This pane evaluates the cumulative Uniform. 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 2,0.*

**Random Numbers**

To generate random numbers from the selected distribution, use the save button on the analysis toolbar.

Random numbers to be generated: 100

**The StatAdvisor**

*This pane allows you to specify the number of observations desired in a random sample from the Uniform. 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.*







Засоренная выборка:

**One-Variable Analysis - NORM+UNIF**

Data variable: NORM+UNIF

50 values ranging from 0,510465 to 17,6035

**The StatAdvisor**

*This procedure is designed to summarize a single sample of data. It will calculate various statistics and graphs. Also included in the procedure are confidence intervals and hypothesis tests. Use the Tabular Options and Graphical Options buttons on the analysis toolbar to access these different procedures.*

**Summary Statistics for NORM+UNIF**

|  |  |
| --- | --- |
| Count | 50 |
| Average | 9,59429 |
| Median | 9,84731 |
| Mode |  |
| Geometric mean | 8,42957 |
| 5% Trimmed mean | 9,62287 |
| 5% Winsorized mean | 9,61303 |
| Variance | 15,9399 |
| Standard deviation | 3,99248 |
| Coeff. of variation | 41,6131% |
| Standard error | 0,564622 |
| 5% Winsorized sigma | 4,02303 |
| MAD | 2,50303 |
| Sbi | 4,07968 |
| Minimum | 0,510465 |
| Maximum | 17,6035 |
| Range | 17,093 |
| Lower quartile | 6,91197 |
| Upper quartile | 11,5486 |
| Interquartile range | 4,6366 |
| 1/6 sextile | 5,95813 |
| 5/6 sextile | 13,9459 |
| Intersextile range | 7,98781 |
| Skewness | -0,029843 |
| Stnd. skewness | -0,0861494 |
| Kurtosis | -0,329264 |
| Stnd. kurtosis | -0,475252 |
| Sum | 479,715 |
| Sum of squares | 5383,58 |

**The StatAdvisor**

*This table shows summary statistics for NORM+UNIF. It includes measures of central tendency, measures of variability, and measures of shape. Of particular interest here are the standardized skewness and standardized kurtosis, which can be used to determine whether the sample comes from a normal distribution. Values of these statistics outside the range of -2 to +2 indicate significant departures from normality, which would tend to invalidate any statistical test regarding the standard deviation. In this case, the standardized skewness value is within the range expected for data from a normal distribution. The standardized kurtosis value is within the range expected for data from a normal distribution.*

**Hypothesis Tests for NORM+UNIF**

Sample mean = 8,91086

Sample median = 8,63448

Sample standard deviation = 4,20054

t-test

Null hypothesis: mean = 9,0

Alternative: not equal

Computed t statistic = -0,212213

P-Value = 0,832377

Do not reject the null hypothesis for alpha = 0,05.

sign test

Null hypothesis: median = 9,0

Alternative: not equal

Number of values below hypothesized median: 54

Number of values above hypothesized median: 46

Large sample test statistic = 0,7 (continuity correction applied)

P-Value = 0,483925

Do not reject the null hypothesis for alpha = 0,05.

signed rank test

Null hypothesis: median = 9,0

Alternative: not equal

Average rank of values below hypothesized median: 48,4815

Average rank of values above hypothesized median: 52,8696

Large sample test statistic = 0,318045 (continuity correction applied)

P-Value = 0,750447

Do not reject the null hypothesis for alpha = 0,05.

chi-squared test

Null hypothesis: sigma = 4,13677

Alternative: not equal

Computed chi-squared statistic = 102,076

P-Value = 0,792047

Do not reject the null hypothesis for alpha = 0,05.

**The StatAdvisor**

*This pane displays the results of tests concerning the population from which the sample of NORM+UNIF comes. The t-test tests the null hypothesis that the mean NORM+UNIF equals 9,0 versus the alternative hypothesis that the mean NORM+UNIF is not equal to 9,0. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The sign test tests the null hypothesis that the median NORM+UNIF equals 9,0 versus the alternative hypothesis that the median NORM+UNIF is not equal to 9,0. It is based on counting the number of values above and below the hypothesized median. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The signed rank test tests the null hypothesis that the median NORM+UNIF equals 9,0 versus the alternative hypothesis that the median NORM+UNIF is not equal to 9,0. It is based on comparing the average ranks of values above and below the hypothesized median. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level. The sign and signed rank tests are less sensitive to the presence of outliers but are somewhat less powerful than the t-test if the data all come from a single normal distribution.*

*The chi-squared test tests the null hypothesis that the standard deviation of NORM+UNIF equals 4,13677 versus the alternative hypothesis that the standard deviation of NORM+UNIF is not equal to 4,13677. Since the P-value for this test is greater than or equal to 0,05, we cannot reject the null hypothesis at the 95,0% confidence level.*







Проверка гипотезы о равенстве среднеквадратичных отклонений нормальной выборки и засоренной:

**Hypothesis Tests**

Sample standard deviation = 4,13677

Sample size = 100

95,0% confidence interval for sigma: [3,63211;4,80558]

Null Hypothesis: standard deviation = 4,24

Alternative: not equal

Computed chi-squared statistic = 94,238

P-Value = 0,766918

Do not reject the null hypothesis for alpha = 0,05.

**The StatAdvisor**

*This analysis shows the results of performing a hypothesis test concerning the standard deviation (sigma) of a normal distribution. The two hypotheses to be tested are:*

*Null hypothesis: sigma = 4,24*

*Alternative hypothesis: sigma <> 4,24*

*Given a sample of 100 observations with a standard deviation of 4,13677, the computed chi-squared statistic equals 94,238. Since the P-value for the test is greater than or equal to 0,05, the null hypothesis cannot be rejected at the 95,0% confidence level. The confidence interval shows that the values of sigma supported by the data fall between 3,63211 and 4,80558.*

**

Вторая нормальная выборка:

**One-Variable Analysis - NORM1**

Data variable: NORM1 (Random Numbers for Dist. 1)

100 values ranging from 1,0418 to 24,0821

**The StatAdvisor**

*This procedure is designed to summarize a single sample of data. It will calculate various statistics and graphs. Also included in the procedure are confidence intervals and hypothesis tests. Use the Tabular Options and Graphical Options buttons on the analysis toolbar to access these different procedures.*

**Summary Statistics for NORM1**

|  |  |
| --- | --- |
| Count | 100 |
| Average | 10,5804 |
| Median | 10,8522 |
| Mode |  |
| Geometric mean | 9,14085 |
| 5% Trimmed mean | 10,4615 |
| 5% Winsorized mean | 10,4765 |
| Variance | 23,1008 |
| Standard deviation | 4,80633 |
| Coeff. of variation | 45,4267% |
| Standard error | 0,480633 |
| 5% Winsorized sigma | 4,67285 |
| MAD | 2,84014 |
| Sbi | 4,74281 |
| Minimum | 1,0418 |
| Maximum | 24,0821 |
| Range | 23,0403 |
| Lower quartile | 7,98972 |
| Upper quartile | 12,949 |
| Interquartile range | 4,95923 |
| 1/6 sextile | 6,23383 |
| 5/6 sextile | 15,1744 |
| Intersextile range | 8,94057 |
| Skewness | 0,271184 |
| Stnd. skewness | 1,1071 |
| Kurtosis | 0,386191 |
| Stnd. kurtosis | 0,78831 |
| Sum | 1058,04 |
| Sum of squares | 13481,5 |

**The StatAdvisor**

*This table shows summary statistics for NORM1. It includes measures of central tendency, measures of variability, and measures of shape. Of particular interest here are the standardized skewness and standardized kurtosis, which can be used to determine whether the sample comes from a normal distribution. Values of these statistics outside the range of -2 to +2 indicate significant departures from normality, which would tend to invalidate any statistical test regarding the standard deviation. In this case, the standardized skewness value is within the range expected for data from a normal distribution. The standardized kurtosis value is within the range expected for data from a normal distribution.*

**Hypothesis Tests for NORM1**

Sample mean = 10,5804

Sample median = 10,8522

Sample standard deviation = 4,80633

t-test

Null hypothesis: mean = 0,0

Alternative: not equal

Computed t statistic = 22,0135

P-Value = 0,0

Reject the null hypothesis for alpha = 0,05.

sign test

Null hypothesis: median = 0,0

Alternative: not equal

Number of values below hypothesized median: 0

Number of values above hypothesized median: 100

Large sample test statistic = 9,9 (continuity correction applied)

P-Value = 0,0

Reject the null hypothesis for alpha = 0,05.

signed rank test

Null hypothesis: median = 0,0

Alternative: not equal

Average rank of values below hypothesized median: 0,0

Average rank of values above hypothesized median: 50,5

Large sample test statistic = 8,68005 (continuity correction applied)

P-Value = 0,0

Reject the null hypothesis for alpha = 0,05.

chi-squared test

Null hypothesis: sigma = 1,0

Alternative: not equal

Computed chi-squared statistic = 2286,98

P-Value = 0,0

Reject the null hypothesis for alpha = 0,05.

**The StatAdvisor**

*This pane displays the results of tests concerning the population from which the sample of NORM1 comes. The t-test tests the null hypothesis that the mean NORM1 equals 0,0 versus the alternative hypothesis that the mean NORM1 is not equal to 0,0. Since the P-value for this test is less than 0,05, we can reject the null hypothesis at the 95,0% confidence level. The sign test tests the null hypothesis that the median NORM1 equals 0,0 versus the alternative hypothesis that the median NORM1 is not equal to 0,0. It is based on counting the number of values above and below the hypothesized median. Since the P-value for this test is less than 0,05, we can reject the null hypothesis at the 95,0% confidence level. The signed rank test tests the null hypothesis that the median NORM1 equals 0,0 versus the alternative hypothesis that the median NORM1 is not equal to 0,0. It is based on comparing the average ranks of values above and below the hypothesized median. Since the P-value for this test is less than 0,05, we can reject the null hypothesis at the 95,0% confidence level. The sign and signed rank tests are less sensitive to the presence of outliers but are somewhat less powerful than the t-test if the data all come from a single normal distribution.*

*The chi-squared test tests the null hypothesis that the standard deviation of NORM1 equals 1,0 versus the alternative hypothesis that the standard deviation of NORM1 is not equal to 1,0. Since the P-value for this test is less than 0,05, we can reject the null hypothesis at the 95,0% confidence level.*







Проверка гипотезы о равенстве мат.ожиданий нормальных выборок:

**Hypothesis Tests for NORM**

Sample mean = 9,33906

Sample median = 9,77057

Sample standard deviation = 3,99259

**Hypothesis Tests for NORM1**

Sample mean = 10,5804

Sample median = 10,8522

Sample standard deviation = 4,80633

**Hypothesis Tests**

Sample means = 9,33906 and 10,5804

Sample standard deviations = 3,99259 and 4,80633

Sample sizes = 50 and 100

95,0% confidence interval for difference between means: -1,24134 +/- 1,5584 [-2,79974;0,317055]

Null Hypothesis: difference between means = 0,0

Alternative: not equal

Computed t statistic = -1,57408

P-Value = 0,117603

Do not reject the null hypothesis for alpha = 0,05.

(Equal variances assumed).

**The StatAdvisor**

*This analysis shows the results of performing a hypothesis test concerning the difference between the means (mu1-mu2) of two samples from normal distributions. The two hypotheses to be tested are:*

*Null hypothesis: mu1-mu2 = 0,0*

*Alternative hypothesis: mu1-mu2 <> 0,0*

*Given one sample of 50 observations with a mean of 9,33906 and a standard deviation of 3,99259 and a second sample of 100 observations with a mean of 10,5804 and a standard deviation of 4,80633, the computed t statistic equals -1,57408. Since the P-value for the test is greater than or equal to 0,05, the null hypothesis cannot be rejected at the 95,0% confidence level. The confidence interval shows that the values of mu1-mu2 supported by the data fall between -2,79974 and 0,317055.*

**

Проверка гипотезы о равенстве среднеквадратичных отклонений нормальных выборок:

**Hypothesis Tests**

Sample standard deviations = 3,99259 and 4,80633

Sample sizes = 50 and 100

95,0% confidence interval for ratio of variances: [0,431989;1,14811]

Null Hypothesis: ratio of variances = 1,0

Alternative: not equal

Computed F statistic = 0,690053

P-Value = 0,151407

Do not reject the null hypothesis for alpha = 0,05.

**The StatAdvisor**

*This analysis shows the results of performing a hypothesis test concerning the ratio of the standard deviations (sigma1/sigma2) of two samples from normal distributions. The two hypotheses to be tested are:*

*Null hypothesis: sigma1/sigma2 = 1,0*

*Alternative hypothesis: sigma1/sigma2 <> 1,0*

*Given one sample of 50 observations with a standard deviation of 3,99259 and a second sample of 100 observations with a standard deviation of 4,80633, the computed F statistic equals 0,690053. Since the P-value for the test is greater than or equal to 0,05, the null hypothesis cannot be rejected at the 95,0% confidence level. The confidence interval shows that the values of var1/var2 supported by the data fall between 0,431989 and 1,14811.*

**Решение в пакете MathCAD**







































Проверки гипотезы о матожидании при известной дисперсии:



















т.к. xleft < zb < xright - гипотеза принимается

Проверка гипотезы о значении матожидания при неизвестной дисперсии:













т.к. zb < xright - гипотеза принимается

Гипотеза о числовом значении дисперсии:









т.к. xleft < zb - гипотеза принимается

Гипотеза о равентсве матожиданий нормальных выборок с известными дисперсиями:













т.к. zb < xleft - гипотеза отвергается

Гипотеза о равенстве матожиданий выборок с неизвестными равными дисперсиями













тк zb < xright - гипотеза принимается

Гипотеза о равенстве дисперсий двух нормальных выборок:













тк xleft < zb < xright - гипотеза принимается