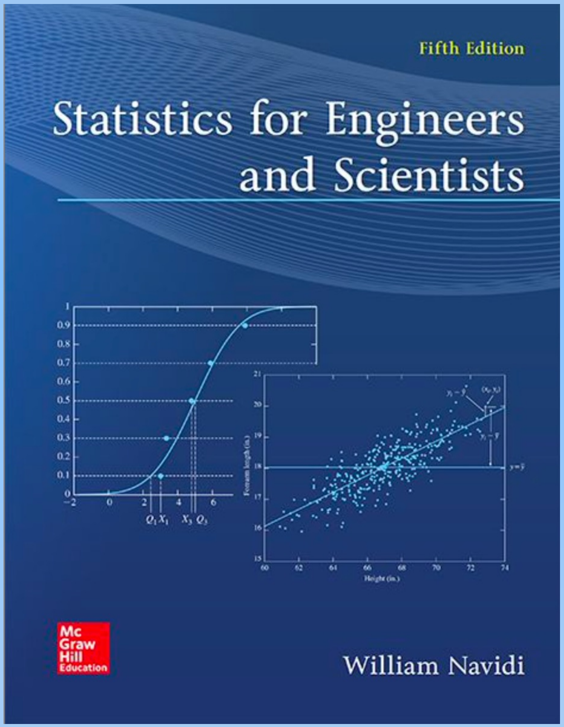


CHAPTER 6

Hypothesis Testing (part 2)

Ch. 6 - Hypothesis Testing

6-1 Large-Sample Tests for a Population Mean

6-2 Drawing Conclusions from the Results of Hypothesis Tests

6-3 **Tests for a Population Proportion**

6-4 **Small-Sample Tests for a Population Mean**

6-5 Large-Sample Tests for the Difference Between Two Means

6-7 Small-Sample Tests for the Difference Between Two Means

6-13 Power

Tests for a Population Proportion

- Many hypothesis-testing situations involve proportions.
- A proportion is the same as a percentage of the population (recall **Chapter 5**).
- For example:

35% of Americans go out for dinner once a week.

z Test for a Population Proportion

Formula for the z Test for Proportions

$$\boxed{z = \frac{\hat{p} - p}{\sqrt{pq/n}}}$$

where $\hat{p} = \dfrac{X}{n}$ sample proportion

$p =$ population proportion

$n =$ sample size

<u>Assumptions for Testing a Proportion:</u>

1. The sample is a random sample.

2. The conditions for binomial experiment are satisfied.

3. The sample size must be sufficiently large
   (np > 10 and nq > 10).

z Test for a Population Proportion

Summary

Let X be the number of successes in n independent Bernoulli trials, each with success probability p; in other words, let $X \sim \text{Bin}(n, p)$.

To test a null hypothesis of the form $H_0: p \le p_0$, $H_0: p \ge p_0$, or $H_0: p = p_0$, assuming that both np_0 and $n(1 - p_0)$ are greater than 10:

- Compute the z-score: $\boxed{z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}.}$
- Compute the P-value. The P-value is an area under the normal curve, which depends on the alternate hypothesis as follows:

Alternate Hypothesis	***P*-value**
$H_1: p > p_0$	Area to the right of z
$H_1: p < p_0$	Area to the left of z
$H_1: p \ne p_0$	Sum of the areas in the tails cut off by z and $-z$

p.424 (textbook)

Example 6.6 (p.423)

In courses on surveying, field work is an important part of the curriculum.

The article “Enhancing Civil Engineering Surveying Learning through Workshops” (J. Awange, A. Faisal Anwar, et al., Journal of Surveying Engineering, 2017, online) reports that in a sample of 67 students studying surveying, 45 said that field work improved their ability to handle unforeseen problems.

Can we conclude that more than 65% of students find that field work improves their ability to handle unforeseen problems?

Example 6.6 (p.423) - SOLUTION

Let p denote the probability that a randomly chosen student believes that field work improves one's ability to handle unforeseen problems.

The null and alternate hypotheses are

- $H_0: p \leq 0.65$
- $H_1: p > 0.65$ (claim)

Example 6.6 (p.423) - SOLUTION

The sample proportion is

$$\hat{p} = \frac{x}{n} = 45/67 = 0.6716$$

Under the null hypothesis, $\hat{p}$ is normally distributed with mean 0.65 and standard deviation

$$\sqrt{\frac{pq}{n}} = \sqrt{\frac{0.65(1-0.65)}{67}} = 0.0583$$

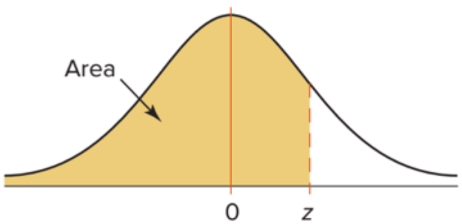
The z score is

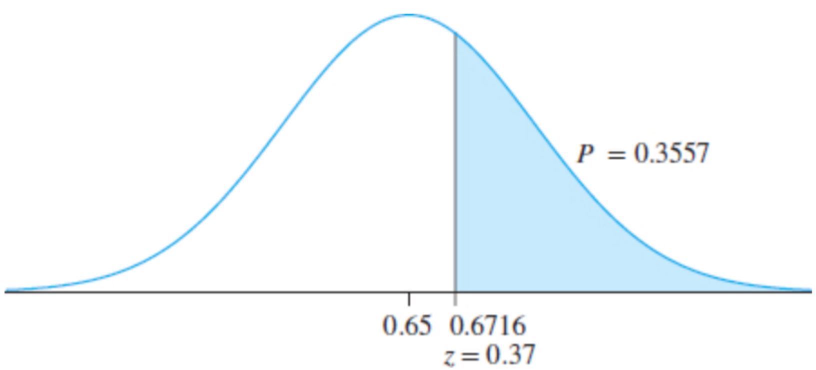
$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} = \frac{0.6716-0.6500}{0.0583} = 0.37$$

Example 6.6 (p.423) - SOLUTION

Cumulative Standard Normal Distribution										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549

The P-value is 1 – 0.6443 = 0.3557





We **do not reject H_0** and therefore **cannot conclude** (which means **cannot support the claim**) that more than 65% of students find that field work improves their ability to handle unforeseen problems.

t Test for a Population Mean

The ***t*-test** is a statistical test for the mean of a population and is used when the population is normally or approximately normally distributed, **σ is unknown**.

The formula for the *t test is*

$$\boxed{t = \frac{\bar{X} - \mu}{s/\sqrt{n}}}$$

The degrees of freedom are **d.f. = *n* – 1**.

Note: When the degrees of freedom are above 30, some textbooks will tell you to use the nearest d.f. value; however, **in this course, you should <u>round down</u> to the nearest d.f. value**. This is a **conservative approach**.

Small-Sample Tests for a Population Mean (p.425)

Summary

Let $X_1, \ldots, X_n$ be a sample from a *normal* population with mean μ and standard deviation σ, where σ is unknown.

To test a null hypothesis of the form $H_0\colon \mu \le \mu_0$, $H_0\colon \mu \ge \mu_0$, or $H_0\colon \mu = \mu_0$:

- Compute the test statistic $\boxed{t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}}$.
- Compute the *P*-value. The *P*-value is an area under the Student's t curve with $n - 1$ degrees of freedom, which depends on the alternate hypothesis as follows:

Alternate Hypothesis	***P*-value**
$H_1\colon \mu > \mu_0$	Area to the right of t
$H_1\colon \mu < \mu_0$	Area to the left of t
$H_1\colon \mu \ne \mu_0$	Sum of the areas in the tails cut off by t and $-t$

- If σ is known, the test statistic is $z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$, and a z test should be performed.

Hypothesis Testing (P-value Method)

All hypothesis-testing situations using the **P-value method** should include the following steps:

Step 1 State the hypotheses and identify the claim.

Step 2 Compute the test value.

Step 3 **Find the P-value.**

Step 4 Make the decision to reject or not reject the null hypothesis.

Step 5 Summarize the results.

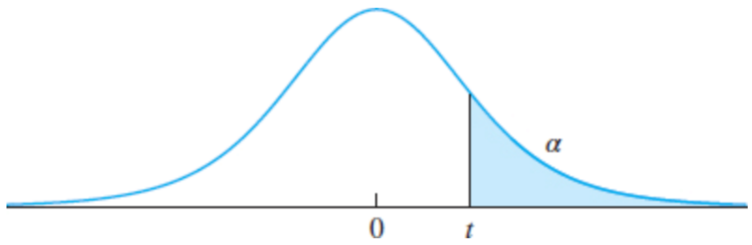
Hypothesis Testing (P-value Method)…

- Typical P-values for the *t* test can be found using Table F.
- However, **not all** P-values for *t* tests can be obtained.
- To find specific P-values for *t* tests, a more precise table, calculator or computer is needed.
- Alternatively, intervals for P-values can be reported.

EXAMPLE

Find the *P*-value when the *t* test value is 2.056, the sample size is 11, and the test is right-tailed.

TABLE A.3 Upper percentage points for the Student's *t* distribution



	α						
v	**0.40**	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947

EXAMPLE

Find the P-value when the t test value is 2.056, the sample size is 11, and the test is right-tailed.

- Hence, the P-value would be contained in the interval $0.025 < \text{P-value} < 0.05$
- If α were **0.1**, you **would reject** the null hypothesis since the P-value must be **less than 0.1**
- But if α were **0.01**, you would **not reject** the null hypothesis since the P-value is **greater than 0.01**
- The P-value obtained from a calculator is **0.033**

Example (p.426)

Spacer collars for a transmission countershaft have a thickness specification of 38.98–39.02 mm. The process that manufactures the collars is supposed to be calibrated so that the mean thickness is 39.00 mm, which is in the center of the specification window.

A sample of six collars is drawn and measured for thickness.

The six thicknesses are 39.030, 38.997, 39.012, 39.008, 39.019, and 39.002. Assume that the population of thicknesses of the collars is approximately normal.

Can we conclude that the process needs recalibration?

Example (p.426) - SOLUTION

Denoting the population mean by μ, the null and alternate hypotheses are

$$\underbrace{H_0 : \mu = 39.00 \quad \text{versus} \quad H_1 : \mu \neq 39.00}$$

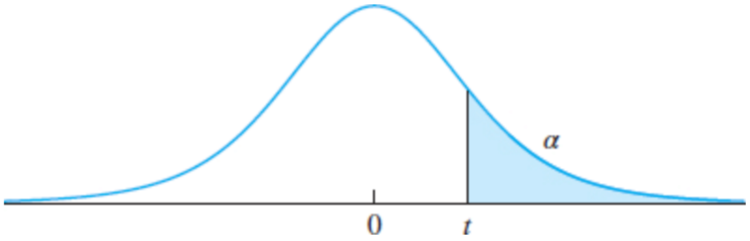
In this example the observed values of the sample mean and standard deviation are $\bar{X} = 39.01133$ and s = 0.011928.

The sample size is n = 6. The null hypothesis specifies that μ = 39.

The value of the test statistic is therefore

$$\boxed{t = \frac{39.01133 - 39.00}{0.011928/\sqrt{6}} = 2.327}$$

TABLE A.3 Upper percentage points for the Student's *t* distribution



v	α = 0.40	0.25	0.10	0.05	0.025	0.01	0.005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947

Two-tailed test: The P-value is between 2*0.025 = 0.05 and 2*0.05 = 0.10 (software gives 0.0675)

Two-tailed test: The P-value is between 2*0.025 = 0.05 and 2*0.05 = 0.10 (software gives 0.0675)

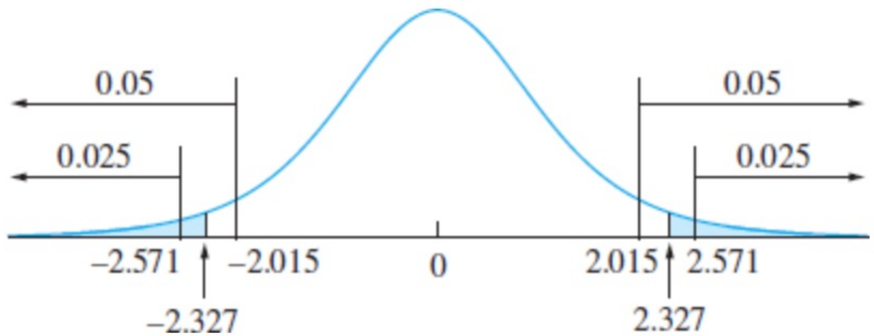


FIGURE 6.7 The null distribution of $t = (\overline{X} - 39.00)/(s/\sqrt{6})$ is Student's t with five degrees of freedom. The observed value of t, corresponding to the observed values $\overline{X} = 39.01133$ and $s = 0.011928$, is 2.327. If H_0 is true, the probability that t takes on a value as extreme as or more extreme than that observed is between 0.05 and 0.10. Because H_0 specified that μ was *equal* to a specific value, both tails of the curve contribute to the P-value.

EXAMPLE Jogger's Oxygen Uptake

A physician claims that joggers' maximal volume oxygen uptake is greater than the average of all adults. A random sample of 15 joggers has a mean of 40.6 milliliters per kilogram (ml/kg) and a standard deviation of 6 ml/kg. If the average of all adults is 36.7 ml/kg, is there enough evidence to support the physician's claim at $\alpha = 0.05$? Assume the variable is normally distributed.

SOLUTION

Step 1 State the hypotheses and identify the claim.

$$H_0\colon \mu = 36.7 \qquad \text{and} \qquad \boxed{H_1\colon \mu > 36.7 \text{ (claim)}}$$

Step 2 Compute the test value. The test value is

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{40.6 - 36.7}{6/\sqrt{15}} = 2.517$$

EXAMPLE **Jogger's Oxygen Uptake**

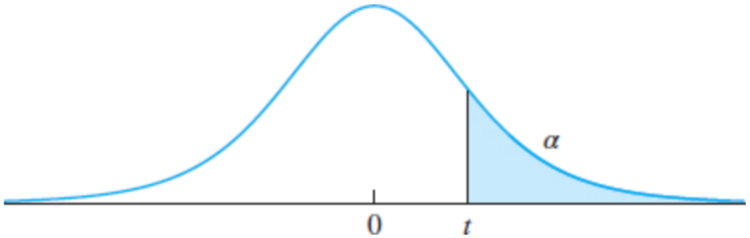
A physician claims that joggers' maximal volume oxygen uptake is greater than the average of all adults. A random sample of 15 joggers has a mean of 40.6 milliliters per kilogram (ml/kg) and a standard deviation of 6 ml/kg. If the average of all adults is 36.7 ml/kg, is there enough evidence to support the physician's claim at $\alpha = 0.05$? Assume the variable is normally distributed.

Step 3 Find the *P*-value. Looking across the row with d.f. = 14 in Table F, you see that 2.517 falls between 2.145 and 2.624, corresponding to $\alpha = 0.025$ and $\alpha = 0.01$ since this is a right-tailed test. Hence, *P*-value > 0.01 and *P*-value < 0.025, or 0.01 < *P*-value < 0.025. That is, the *P*-value is somewhere between 0.01 and 0.025. (The *P*-value obtained from a calculator is 0.012.)

Step 4 Reject the null hypothesis since *P*-value < 0.05 (that is, *P*-value < α).

Step 5 There is enough evidence to support the claim that the joggers' maximal volume oxygen uptake is greater than 36.7 ml/kg.

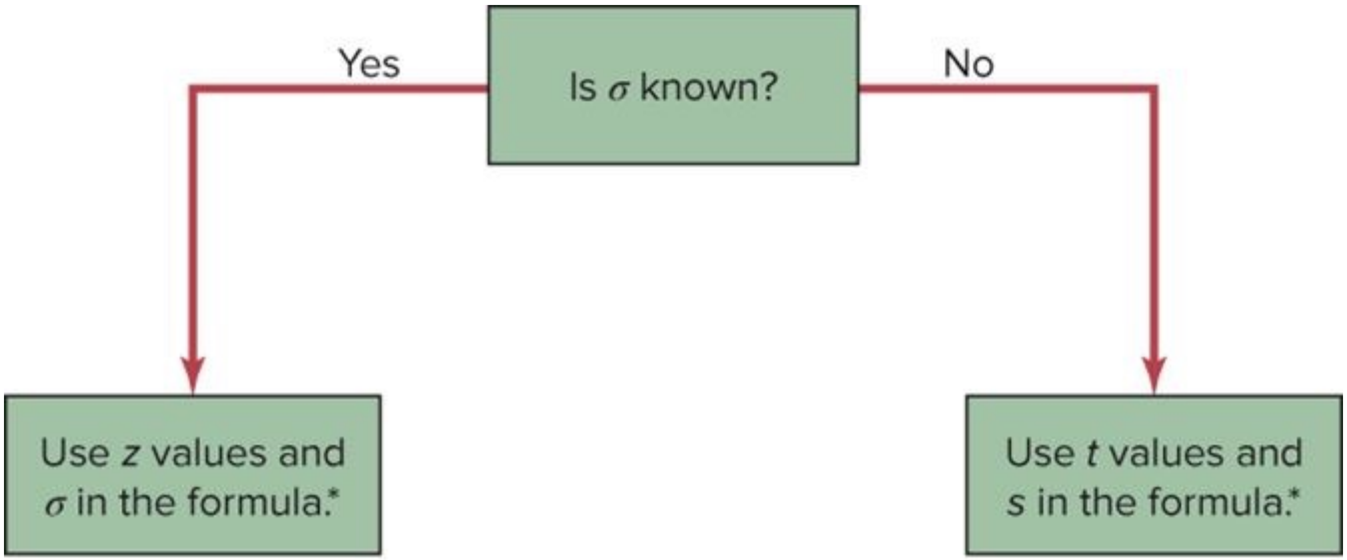
TABLE A.3 Upper percentage points for the Student's *t* distribution



v	α = 0.40	0.25	0.10	0.05	0.025	0.01	0.005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947

Right-tailed test: The P-value is between 0.01 and 0.025

Using z Test or *t* Test?



*If $n < 30$, the variable must be normally distributed.

Example 6.8

At the beginning of this section, we described a sample of six spacer collars, whose thicknesses (in mm) were 39.030, 38.997, 39.012, 39.008, 39.019, and 39.002. We denoted the population mean thickness by μ and tested the hypotheses

$$\boxed{H_0: \mu = 39.00 \quad \text{versus} \quad H_1: \mu \neq 39.00}$$

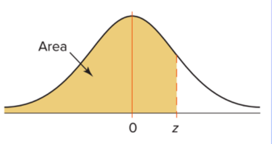
Now assume that these six spacer collars were manufactured just after the machine that produces them had been moved to a new location. Assume that on the basis of a very large number of collars manufactured before the move, the population of collar thicknesses is known to be very close to normal, with standard deviation $\sigma = 0.010$ mm, and that it is reasonable to assume that the move has not changed this. On the basis of the given data, can we reject H_0?

$$z = \frac{39.01133 - 39.000}{0.010/\sqrt{6}} = 2.78 \qquad \cancel{t = \frac{39.01133 - 39.00}{0.011928/\sqrt{6}} = 2.327}$$

The P-value is **0.0054**, so **H_0 can be rejected**.

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986



Two-tailed test: The P-value = (1-0.9973)*2 = 0.0054