Advanced Statistics

Dr. Syed Faisal Bukhari
Associate Professor
Department of Data Science
Faculty of Computing and Information Technology
University of the Punjab

Textbooks

- **Probability & Statistics for Engineers & Scientists**, Ninth Edition, Ronald E. Walpole, Raymond H. Myer

- **Elementary Statistics: Picturing the World,** 6th Edition, Ron Larson and Betsy Farber

- **Elementary Statistics**, 13th Edition, Mario F. Triola

Reference books

- **Probability and Statistical Inference, Ninth Edition,** Robert V. Hogg, Elliot A. Tanis, Dale L. Zimmerman
- **Probability Demystified**, Allan G. Bluman
- **Practical Statistics for Data Scientists: 50 Essential Concepts,** Peter Bruce and Andrew Bruce
- **Schaum's Outline of Probability,** Second Edition, Seymour Lipschutz, Marc Lipson
- **Python for Probability, Statistics, and Machine Learning,** José Unpingco

References

- **Elementary Statistics**, 13th Edition, Mario F. Triola

These notes contain material from the above resource.

- https://hbr.org/2016/02/a-refresher-on-statistical-significance

These notes contain material from the above resources.

What Is Statistical Significance?

“Statistical significance helps quantify whether a result is likely due to chance or to some factor of interest, ”says Redman.

When a **finding is significant**, it simply means **you can feel confident that’s it real**, not that you just got lucky (or unlucky) in choosing the sample.

Nonparametric (or Distribution-Free) Tests

Parametric tests have requirements about the distribution of the populations involved; **nonparametric (or distribution-free) tests** do not require **that samples come from populations with normal distributions** or **any other particular distributions**.

Misleading Terminology

The term ***distribution-free test*** correctly indicates that a test does not **require a particular distribution**. The term ***nonparametric tests* is misleading** in the sense that it suggests that the tests are **not based on a parameter**, but there **are some nonparametric tests** that are **based on a parameter such as the median**.

Due to the widespread use of the **term *nonparametric test*,** we use that terminology, but we define it to be **a test that does not require a particular distribution**.

Advantages and Disadvantages

Advantages of Nonparametric Tests

1. Because **nonparametric tests** have **less rigid requirements** than **parametric tests**, they can be **applied to a wider variety** of situations.

2. **Nonparametric tests** can be applied to **more data types than parametric tests**. For example, **nonparametric tests** can be used with data **consisting of ranks**, and they can be used with **categorical data**, such as **genders of survey respondents**.

Disadvantages of Nonparametric Tests

1. Nonparametric tests tend to **waste information** because **exact numerical data** are often **reduced to a qualitative form**. For example, with the **nonparametric sign test** weight losses by **dieters** are recorded simply as **negative signs**, and the **actual magnitudes** of the **weight losses are ignored**.

2. **Nonparametric tests** are **not as *efficient* as parametric tests**, so a **nonparametric test generally** needs **stronger evidence** (such as a larger sample or greater differences) in order to **reject a null hypothesis**.

Efficiency: Comparison of Parametric and Nonparametric Tests

Table on the next slide shows that **several nonparametric tests have efficiency ratings above 0.90**, so the **lower efficiency might not** be an important factor in choosing **between parametric and nonparametric tests**.

However, because **parametric tests do have higher efficiency** ratings than their **nonparametric counterparts**, it's generally better to use the parametric tests **when their required assumptions are satisfied**.

Efficiency: Comparison of Parametric and Nonparametric Tests

Application	Parametric Test	Nonparametric Test	Efficiency Rating of Nonparametric Test with Normal Populations
Matched pairs of sample data	*t* test	Sign test or	0.63
		Wilcoxon signed-ranks test	0.95
Two independent samples	*t* test	Wilcoxon rank-sum test	0.95
Three or more independent samples	Analysis of variance (*F* test)	Kruskal-Wallis test	0.95
Correlation	Linear correlation	Rank correlation test	0.91
Randomness	No parametric test	Runs test	No basis for comparison

Rank

DEFINITION

Data are *sorted* when they are arranged according to some criterion, such as **smallest to largest** or **best to worst**. A **rank** is a **number assigned** to an **individual sample** item according to its **order in the sorted list**. The first item is assigned a rank of 1, the second item is assigned a rank of 2, and so on.

Handling Ties Among Ranks

If a **tie in ranks occurs**, one very common procedure is to **find the mean of the ranks** involved in the **tie** and then **assign this mean rank** to each of the **tied items**, as in the following example.

Handling Ties Among Ranks

EXAMPLE The golf scores (for one hole) of 4, 5, 5, 5, 10, 11, 12, and 12 are given ranks of 1, 3, 3, 3, 5, 6, 7.5, and 7.5, respectively. The table below illustrates the procedure for handling ties.

Sorted Data	Preliminary Ranking	Rank
4	1	**1**
5 5 5	2 3 Mean is 3. 4	**3** **3** **3**
10	5	**5**
11	6	**6**
12 12	7 8 Mean is 7.5.	**7.5** **7.5**

Sign Test

DEFINITION

The **sign test** is a nonparametric (distribution-free) test that uses **positive** and **negative signs** to test different claims, including these:

1. Claims involving **matched pairs of sample data**

2. Claims involving **nominal data** with **two categories**

3. Claims about the **median** of a **single population**

Basic Concept of the Sign Test

The basic idea underlying the **sign test** is to **analyze the frequencies of positive and negative signs** to determine whether they are **significantly different**.

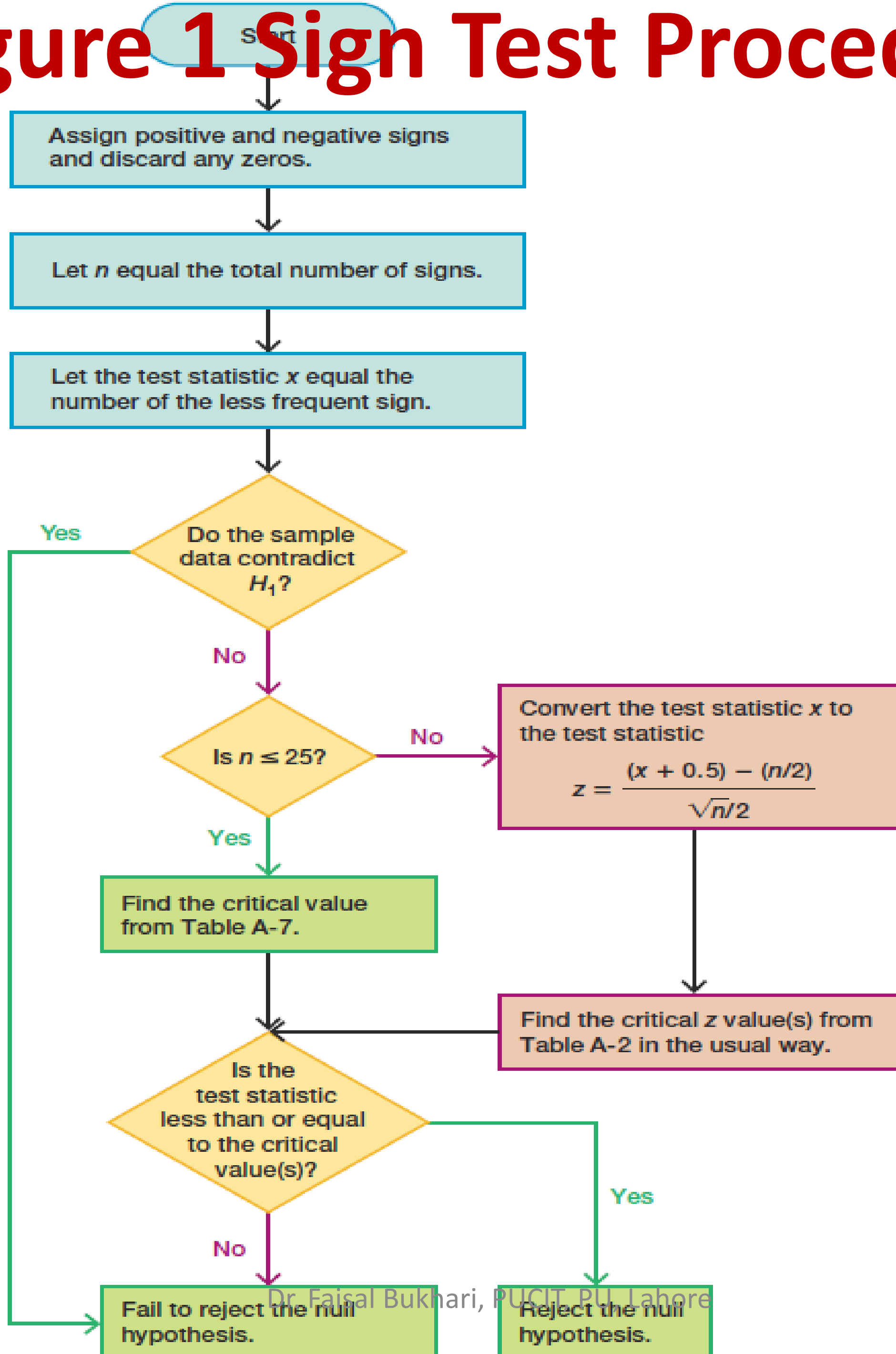
For example, consider the results of clinical trials of the XSORT

method of gender selection. Among **726 couples** who used the XSORT method in trying to have a baby girl, **668 couples did have baby girls. Is 668 girls** in **726 births significant**? Common sense should suggest that **668 girls in 726 births is significant**, but what about **365 girls in 726 births**? Or **400 girls in 726 births**?

The **sign test allows** us to determine when such **results are significant**.

For **consistency and simplicity**, we will use a **test statistic based** on the number of times that the ***less frequent* sign occurs**.

Figure 1 Sign Test Procedure



Sign Test

Objective

Use **positive** and **negative signs** to test a claim falling into **one of** the following **three categories**:

1. Matched Pairs

Subtract the **second value** in each pair from **the first**, record the sign of the difference, and **ignore any 0s**.

2. Nominal Data with Two Categories

Represent each member of one **category by a positive sign** and represent each member of the other **category by a negative sign**.

3. Median of a Single Population

Subtract the **median from each sample value**, record the sign of the difference, and **ignore any 0s**.

Notation

x = the number of **times the less frequent sign** occurs

n = the total number of **positive and negative signs** combined

Requirements

The sample data are a simple random sample.

Note: There is ***no* requirement** that the **sample data** come from a **population with a particular distribution**, such as a normal distribution.

Test Statistic

If $n \leq 25$: Test statistic is ***x* = the number of times the less frequent sign occurs.**

If $n > 25$: Test statistic is

$$Z = \frac{(x + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$$

***P*-Values**

P-values are often provided by technology, or *P*-values can often be found using the *z* test statistic.

Critical Values

1. If $n \leq 25$, critical ***x* values** are found in Table A-7.

2. If $n > 25$, critical ***z* values** are found in Table A-2.

Hint: Because *x* or *z* is based on the ***less* frequent sign**, all **one-sided tests are treated** as if they were **left-tailed tests**.

EXAMPLE Data Contradicting the Alternative Hypothesis

Among 945 couples who used the XSORT method of gender selection, **66 had boys**, so the sample proportion of boys is $\frac{66}{945}$, **or 0.0698** (based on data from the Genetics & IVF Institute).

Consider the claim that the XSORT method of gender selection *increases* the likelihood of baby *boys* so that the probability of a boy is $p > 0.5$. This claim of $p > 0.5$ becomes the alternative hypothesis.

Using common sense, we see that with a sample proportion of boys of **0.0698**, we can **never support a claim that $p > 0.5$**. (We would need a sample proportion of boys *greater* than 0.5 by a significant amount.) Here, the sample proportion of $\frac{66}{945}$, **or 0.0698**, ***contradicts* the alternative hypothesis** because it is **not greater than 0.5.**

INTERPRETATION

An **alternative hypothesis** can never be **supported with data that contradict it**. The sign test will show that 66 boys in 945 births is significant, but it is significant in the wrong direction.

We can never support a claim that $\boldsymbol{p > 0.5}$ with a sample proportion of $\frac{66}{945}$, **or 0.0698**, which is *less than* 0.5.

Claims About Matched Pairs

When using the sign test with data that are matched pairs, we convert the raw data to **positive and negative signs** as follows:

1. Subtract each value of the **second variable** from the **corresponding value** of the **first variable**.

2. Record only the ***sign* of the difference found** in Step 1. **Exclude *ties*** by **deleting any matched pairs** in which both values are equal.

The main concept underlying this use of the sign test is as follows:

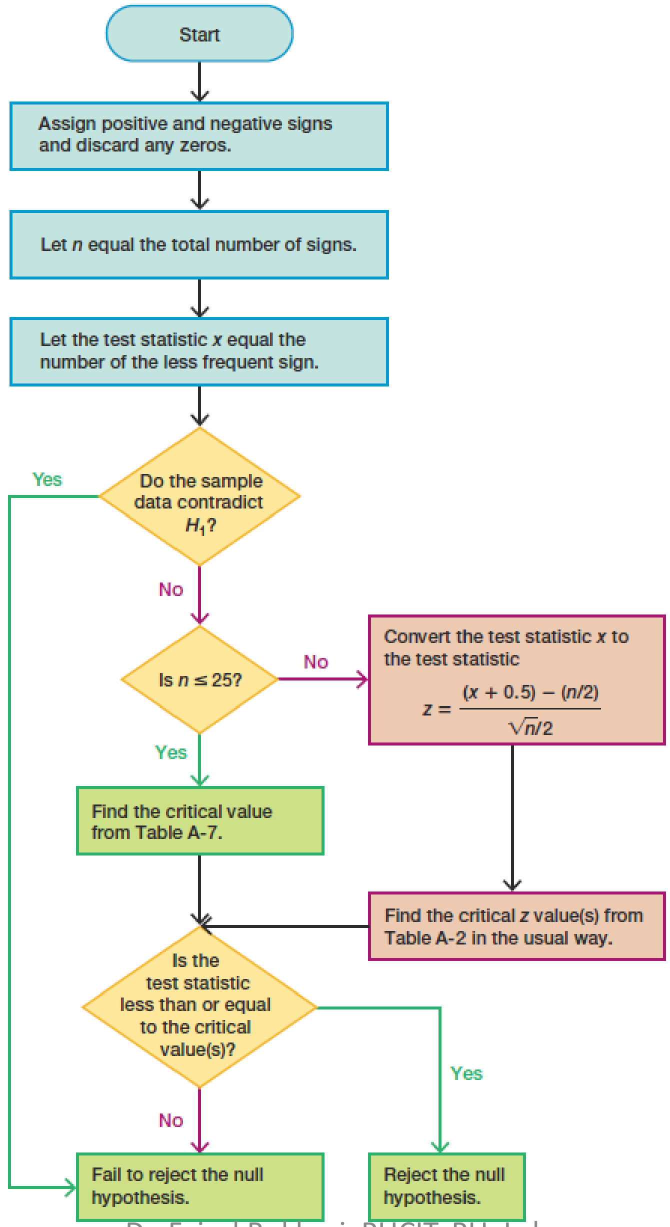
If the two sets of data have equal medians, the number of positive signs should be approximately equal to the number of negative signs.

Data Set: Measured and Reported

Measured weights and heights matched with the weights and heights that were reported when **5755 subjects** (first five rows shown here) aged 16 and over were asked for those values. Values are given in both Imperial and metric systems. Weights are given in pounds and

GENDER (1 = M)	MEASURED WEIGHT (LB)	REPORTED WEIGHT (LB)	MEASURED HEIGHT (IN)	REPORTED HEIGHT (IN)	MEASURED WEIGHT (KG)	REPORTED WEIGHT (KG)	MEASURED HEIGHT (CM)	REPORTED HEIGHT (CM)
1	209.0	212	72.6	74	94.8	96.2	184.5	188.0
1	199.3	193	67.5	68	90.4	87.5	171.4	172.7
1	183.9	182	67.0	69	83.4	82.6	170.1	175.3
0	242.1	220	63.3	64	109.8	99.8	160.9	162.6
0	121.7	125	64.9	64	55.2	56.7	164.9	162.6

Figure 1 Sign Test Procedure



Is There a Difference Between Measured and Reported Weights?

EXAMPLE

Listed below are measured and reported weights (lb) of random male subjects (from Data Set **"Measured and Reported"**). Use a 0.05 significance level to test the claim that for males, the differences "measured weight–reported weight" have a median equal to 0.

Table 1 Measured and Reported Male Weights

Measured	220.0	268.7	213.4	201.3	107.1	172.0	187.4	132.5	122.1	151.9
Reported	220	267	210	204	107	176	187	135	122	150
Sign of Difference	0	+	+	−	+	−	+	−	+	+

REQUIREMENT CHECK The only requirement of the sign test is that the sample data are a simple random sample, and that requirement is satisfied.

- If there is **no difference between measured weights and reported weights**, the numbers of positive and negative signs should be **approximately equal**.

- In Table 1 we have **6 positive signs**, **3 negative signs**, and **1 difference of 0.**

- We **discard the difference of 0** and proceed using only the 6 positive signs and 3 negative signs.

- The **sign test tells us whether or not the numbers of positive and negative signs are approximately equal**.

$\mathbf{H_0}$: **There is no difference.** (The median of the differences is equal to 0.)

$\mathbf{H_1}$: **There is a difference.** (The median of the differences is not equal to 0.)

Following the sign test procedure summarized in Figure 1, we let ***n* = 9** (the total number of positive and negative signs) and we let ***x* = 3** (the number of the less frequent sign, or the smaller of 3 and 6).

The sample data do not contradict $\mathbf{H_1}$, because there is a difference between the **6 positive signs** and the **3 negative signs**. The sample data **show a difference**, and we need to continue with the **test to determine whether** that **difference is significant**.

- Figure 1 shows that **with *n* = 9**, we should proceed to find the **critical value from Table A-7**. We refer to Table A-7, where the **critical value of 1** is found for ***n* = 9 and α = 0.05** in two tails.

- Since ***n*** $\leq$ **25**, the test statistic is ***x* = 3** (and we do not convert *x* to a *z* score). With a test statistic of *x* = 3 and a critical ***x* value of 1**, we fail to reject the null hypothesis of no difference.

  Test-statistic $\leq$ Critical value

  $\Rightarrow$ **3** $\leq$ **1** (False, fail to reject $\mathbf{H_0}$)

- (See Note 2 included with Table A-7: **"Reject the null hypothesis if the number of the less frequent sign (*x*) is less than or equal to the value in the table."**

$\mathbf{x} \leq \boldsymbol{critical\ value}$

3 ≤ 1 (False)

Because *x* = 3 is *not* less than or equal to the critical value of 1, we fail to reject the null hypothesis.)

There is not sufficient evidence to warrant rejection of the claim that for males, the median of the differences “measured weight – reported weight” is equal to 0.

INTERPRETATION

We conclude that there is not **sufficient evidence to reject** the claim that for males, there is no difference between measured weights and reported weights.

TABLE A-7 Critical Values for the Sign Test

	α			
	.005 (one tail)	.01 (one tail)	.025 (one tail)	.05 (one tail)
n	.01 (two tails)	.02 (two tails)	.05 (two tails)	.10 (two tails)
1	*	*	*	*
2	*	*	*	*
3	*	*	*	*
4	*	*	*	*
5	*	*	*	0
6	*	*	0	0
7	*	0	0	0
8	0	0	0	1
9	0	0	1	1
10	0	0	1	1
11	0	1	1	2
12	1	1	2	2
13	1	1	2	3
14	1	2	2	3
15	2	2	3	3
16	2	2	3	4
17	2	3	4	4
18	3	3	4	5
19	3	4	4	5
20	3	4	5	5
21	4	4	5	6
22	4	5	5	6
23	4	5	6	7
24	5	5	6	7
25	5	6	7	7

TABLE A-7

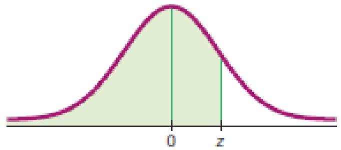
NOTES:

1. * indicates that it is not possible to get a value in the critical region, so fail to reject the null hypothesis.

2. Reject the null hypothesis if the number of the less frequent sign (*x*) is less than or equal to the value in the table.

3. For values of *n* greater than 25, a normal approximation is used with

$$Z = \frac{(x + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$$



POSITIVE *z* Scores

TABLE A-2 *(continued)* Cumulative Area from the LEFT

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
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50 and up	.9999									

NOTE: For values of *z* above 3.49, use 0.9999 for the area.

*Use these common values that result from interpolation:

z Score	Area
1.645	0.9500
2.575	0.9950

Common Critical Values

Confidence Level	Critical Value
0.90	1.645
0.95	1.96
0.99	2.575

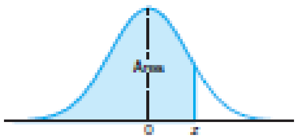


Table A.3 Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table A.3 (continued) Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Claims Involving Nominal Data with Two Categories

We defined **nominal data** to be data that **consist of names, labels, or categories only.**

The nature of nominal data limits the calculations that are possible, but we can identify the ***proportion* of the sample data** that belong to a particular category, and we can test claims about the corresponding population proportion *p*.

The following example uses nominal data consisting of genders (girls/boys). The sign test is used by representing **girls with positive (+)** signs and **boys with negative (-) signs**. (Those signs are chosen arbitrarily—honest.)

EXAMPLE Gender Selection

The Genetics & IVF Institute conducted a clinical trial of its methods for gender selection for babies. Before the clinical trials were concluded, **879 of 945 babies** born to parents using the XSORT method of gender selection were **girls**. Use **the sign test** and a 0.05 significance level to test the claim that **this method of gender selection is effective in increasing the likelihood of a baby girl**.

Solution

REQUIREMENT CHECK The only requirement is that **the sample is a simple random sample**. Based on the design of this experiment, we can assume that the sample data are a simple random sample.

Let ***p* denote the population proportion of baby girls**. The claim that girls are more likely with the XSORT method can be expressed as $p > 0.5$, so the null and alternative hypotheses are as follows:

$\mathbf{H_0}$: $p = 0.5$ (the proportion of girls is equal to 0.5)

$\mathbf{H_1}$: $p > 0.5$ (girls are more likely)

- Denoting **girls by positive signs (+)** and **boys by negative signs (-)**, we have 879 positive signs and 66 negative signs. Using the sign test procedure summarized in Figure 1, we let the **test statistic *x*** be the **smaller of 879 and 66**, so ***x* = 66 boys**.

- Instead of trying to determine **whether 879 girls is high enough** to be significantly high, we proceed with the equivalent goal of trying to determine **whether 66 boys is low enough** to be significantly low, so we treat the **test as a left-tailed test**.

- The sample data **do not contradict the alternative hypothesis** because the sample proportion of girls is $\frac{879}{945}$, **or 0.930**, which is **greater than 0.5**, as in the above alternative hypothesis.

Continuing with the procedure in Figure 1, we note that the value of ***n* = 945** is **greater than 25**, so the test statistic ***x* = 66** is converted (using a correction for continuity) to the test statistic *z* as follows:

$$Z = \frac{(x + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$$

$$Z = \frac{(66 + 0.5) - \frac{945}{2}}{\frac{\sqrt{945}}{2}} = -26.41$$

***P*-Value** We could use the test statistic of ***z* = -26.41** to find the left-tailed ***P*-value of 0.0000** (using Table). That **low *P*-value causes us to reject the null hypothesis**.

- ***Reject H_0*** if the ***P*-value** $\le \alpha$ (where α is the significance level, such as 0.05). ***H_0*** is called **statistically significant**.
- ***Fail to reject H_0*** if the ***P*-value** > α. ***H_0* is not statistically significant**.

Critical Value With a = 0.05 in a left-tailed test, the critical value is *z* = -1.645.

Figure 1 shows that the **test statistic *z* = -26.41** is in the critical region bounded by ***z* = -1.645**, so we **reject the null hypothesis** that the proportion of girls is equal to 0.5.

There is sufficient sample evidence to support the claim that girls are more likely with the XSORT method.

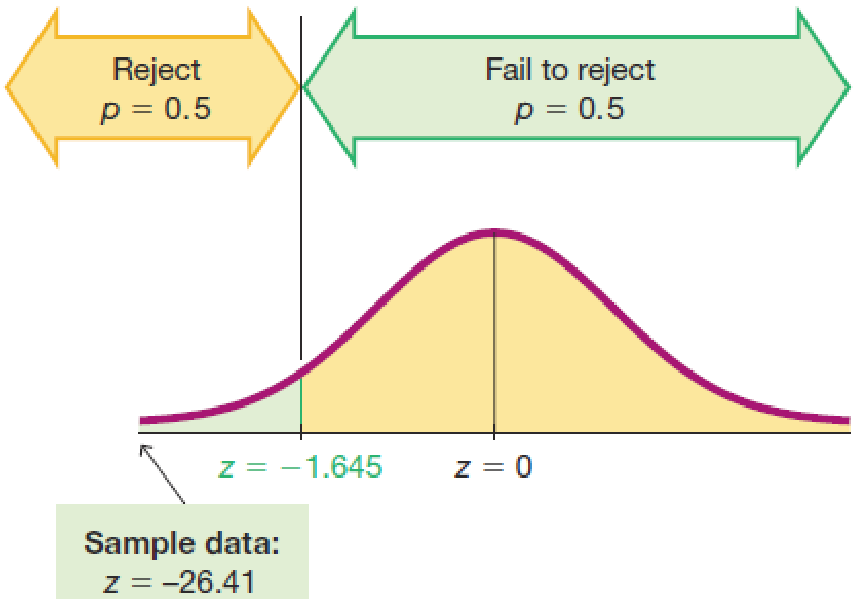


FIGURE-2 Testing Effectiveness of the XSORT Gender Selection Method

INTERPRETATION

The XSORT method of gender selection **does appear to be associated with an increase in the likelihood of a girl**, so this method appears to be effective (but this hypothesis test does not prove that the XSORT method is the *cause* of the increase).

Data Set : Body Temperatures

Body temperatures 1F2 are from 107 subjects taken on two consecutive days at 8 AM and 12 AM (first five rows shown here). **SEX** is gender of subject, and **SMOKE** indicates if subject smokes (Y) or does not smoke (N). Data provided by Dr. Steven Wasserman, Dr. Philip Mackowiak, and Dr. Myron Levine of the University of

Maryland.

SEX	SMOKE	DAY 1—8 AM	DAY 1—12 AM	DAY 2—8 AM	DAY 2—12 AM
M	Y	98.0	98.0	98.0	98.6
M	Y	97.0	97.6	97.4	—
M	Y	98.6	98.8	97.8	98.6
M	N	97.4	98.0	97.0	98.0
M	N	98.2	98.8	97.0	98.0

EXAMPLE Body Temperatures

Data Set "Body Temperatures" in previous slide includes measured body temperatures of adults. Use the 106 temperatures listed for 12 AM on Day 2 with the sign test to test the claim that the **median is less than 98.6F**. Use a 0.05 significance level. Of the 106 subjects, 68 had temperatures below 98.6

REQUIREMENT CHECK The only requirement is that the sample is a simple random sample. Based on the design of this experiment, we assume that the sample data are a simple random sample.

The claim that the median is less than 98.6F is the alternative hypothesis, while the null hypothesis is the claim that the median is equal to 98.6F.

***H_0*: Median is equal to 98.6F.** (median = 98.6F)

***H_1*: Median is less than 98.6F.** (median < 98.6F)

- Following the procedure outlined in Figure 1, we use a negative sign to represent each temperature below 98.6F, and we use a positive sign for each temperature above 98.6F.

- We discard the 15 data values of 98.6, since they result in differences of zero. We have **68 negative signs and 23 positive** signs, **so *n* = 91 and *x* = 23 (the number of the less frequent sign)**. The sample data do not contradict the alternative hypothesis, because most of **the 91 temperatures are below 98.6 F**.

- The value of *n* exceeds 25, so we convert the test statistic *x* to the test statistic *z:*

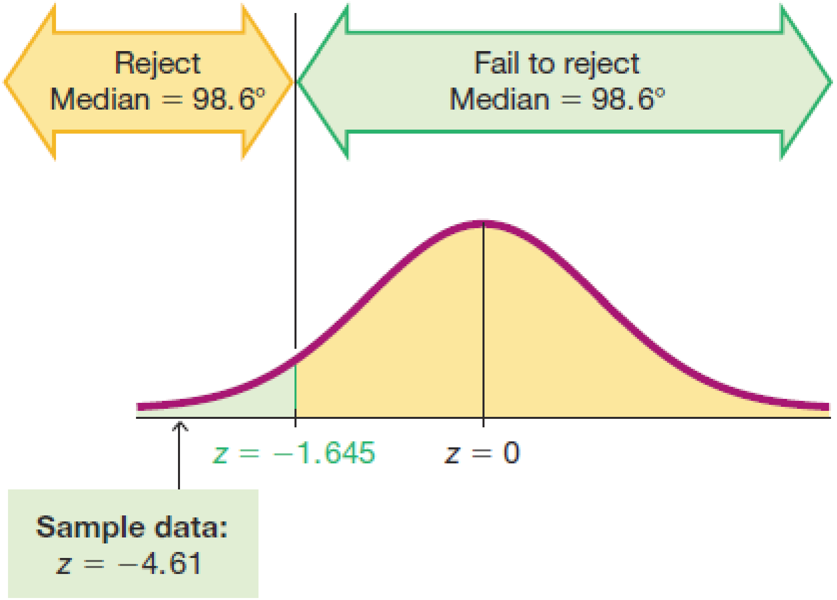
$$Z = \frac{(x + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$$

$$Z = \frac{(23 + 0.5) - \frac{91}{2}}{\frac{\sqrt{91}}{2}} = -4.61$$

In this left-tailed test, the test statistic of $z = -4.61$ yields a *P*-value of 0.0000 (Table: 0.0001). Because that *P*-value is so small, we reject the null hypothesis.

Reject H_0 if the ***P*-value** $\leq \alpha$ (where α is the significance level, such as 0.05). H_0 is called **statistically significant**.

- **Critical Value** In this left-tailed test with a = 0.05, use **Table A-2** to get the **critical *z* value of -1.645**. From Figure 1 on the next page we see that the test statistic of ***z* = -4.61** is within the critical region, **so reject the null hypothesis**.



INTERPRETATION

There is sufficient sample evidence to support the claim that the median body temperature of **healthy adults is less than 98.6F**. It is not equal to 98.6F, as is commonly believed.