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.

Wilcoxon signed-ranks test

DEFINITION

The **Wilcoxon signed-ranks test** is a nonparametric test that uses ranks for these applications:

1. Testing a claim that a **population of matched pairs** has the property that the **matched pairs have differences** with a **median equal to zero.**

2. Testing a claim that a **single population of individual** values has a **median equal** to some **claimed value.**

Wilcoxon Signed-Ranks Test

Objective: Use the Wilcoxon signed-ranks test for the following tests:

Matched Pairs: Test the claim that a population of matched pairs has the property that the matched pairs have differences with a **median equal to zero**.

One Population of Individual Values: Test the claim that a population has a **median equal to some claimed value**. (By pairing each sample value with the claimed median, we again work with matched pairs.)

Wilcoxon Signed-Ranks Test

Notation

***T* =** the smaller of the following two sums:

1. The sum of the positive ranks of the nonzero differences *d*

2. The absolute value of the sum of the negative ranks of the nonzero differences *d*

Requirements

1. The data are a **simple random sample**.

2. The population of **differences has a distribution that is approximately *symmetric*,** meaning that the **left half of its histogram** is roughly a **mirror image of its right half**. (For a sample of matched pairs, obtain differences by subtracting the second value from the first value in each pair; for a sample of individual values, obtain differences by subtracting the value of the claimed median from each sample value.)

Note: There is *no* requirement that the data have a normal distribution.

Test Statistic

If ***n*** $\leq$ **30:** Test statistic is ***x* = the number of times the less frequent sign occurs**.

If *n* $>$ 30, the test statistic is

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

P-Values

P-values are often provided by technology, or *P*-values can be found using the *z* test statistic and Table A-2.

Critical Values

1. If ***n*** $\leq$ **30**, the critical *T* value is found in **Table A-8.**

2. If *n n* $>$ 30, the critical *z* values are found in **Table A-2.**

TABLE A-8 Critical Values of T for the Wilcoxon Signed-Ranks Test

	α			
n	.005 (one tail) .01 (two tails)	.01 (one tail) .02 (two tails)	.025 (one tail) .05 (two tails)	.05 (one tail) .10 (two tails)
5	*	*	*	1
6	*	*	1	2
7	*	0	2	4
8	0	2	4	6
9	2	3	6	8
10	3	5	8	11
11	5	7	11	14
12	7	10	14	17
13	10	13	17	21
14	13	16	21	26
15	16	20	25	30
16	19	24	30	36
17	23	28	35	41
18	28	33	40	47
19	32	38	46	54
20	37	43	52	60
21	43	49	59	68
22	49	56	66	75
23	55	62	73	83
24	61	69	81	92
25	68	77	90	101
26	76	85	98	110
27	84	93	107	120
28	92	102	117	130
29	100	111	127	141
30	109	120	137	152

Table A-8

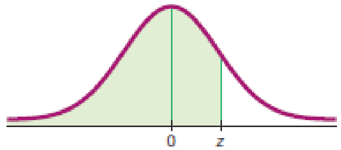
NOTES:

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

2. Conclusions:

Reject the null hypothesis if the **test statistic *T* is less than or equal to the critical value** found in this table.

Fail to reject the null hypothesis if the test statistic *T* is greater than the critical value found in the table.



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

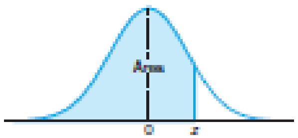


Table A.3 Areas under the Normal Curve

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

Wilcoxon Signed-Ranks Procedure

- The following example includes the **eight steps of the Wilcoxon signed-ranks procedure**. This procedure requires that you **sort data, then assign ranks**. When working with larger **data sets, sorting and ranking become tedious**, but technology can be used to **automate that process**.

- **Stemplots** can also be very helpful in sorting data.

Data Set 4: 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 kilograms. Heights are given in inches and centimeters. Data are from the National Center for Health Statistics.

TI-83/84 list names (MESREPT)*: MRGND, MWTLB, RWLB, MHTIN, RHTIN, MWTKG, RWKG, MHTCM, RHTCM

**NOTE: TI lists are limited to 500 rows due to calculator memory constraints.*

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

EXAMPLE Measured and Reported Weights

The first two rows of Table 1 include measured and reported weights from a simple random sample of eight different male subjects (from Data Set 4 "Measured and Reported"). The data are matched, so each measured weight is paired with the corresponding reported weight. Assume that we want to use the **Wilcoxon signed-ranks test** with a 0.05 significance level to test the claim that there is a **significant difference between measured weights and reported weights of males**. That is, assume that we want to test the null hypothesis that the matched pairs are from a population of matched pairs with differences having a **median equal to zero**.

Table Measured and Reported Weights (kg)

Measured Weights	152.6	149.3	174.8	119.5	194.9	180.3	215.4	239.6
Reported Weights	150	148	170	119	185	180	224	239
d (difference)	2.6	1.3	4.8	0.5	9.9	0.3	−8.6	0.6
Rank of $\|d\|$	5	4	6	2	8	1	7	3
Signed rank	5	4	6	2	8	1	−7	3

SOLUTION

REQUIREMENT CHECK

(1) The data are a **simple random sample**, as required.

(2) The second requirement is that **the population of differences has a distribution that is approximately *symmetric*,** meaning that the left half of its **histogram is roughly a mirror image of its right half**. A histogram of the differences in the third row of Table 1 shows that the difference between the **left and right sides is not too extreme**, so we will consider this requirement to be satisfied.

Wilcoxon Signed-Ranks Procedure

Step 1: For each pair of data, find the **difference *d*** by subtracting the **second value** from the **first value**. Discard any **pairs that have a difference of 0**.

EXAMPLE: The third row of Table 1 lists the differences found by subtracting the reported weights from the measured weights. Ignore any differences equal to 0.

Step 2: ***Ignore the signs of the differences,*** then sort the differences from **lowest to highest and replace the differences** by the **corresponding rank value** (as described previously). When differences have the **same numerical value**, assign to them the **mean of the ranks involved in the tie**.

Wilcoxon Signed-Ranks Procedure

EXAMPLE: The fourth row of Table 1 shows the ranks of the values of d. The **smallest value of d is 0.3**, so it is assigned the rank of 1. The next smallest value of d is 0.5, so it is assigned the **rank of 2**. If there had **been any ties**, they would have been assigned the **mean of the ranks** involved in the tie.

Step 3: Attach to each **rank the sign of the difference** from which it came. That is, insert the signs that **were ignored in Step 2**.

EXAMPLE: The bottom row of Table 1 lists the same ranks found in the fourth row, but the signs of the differences shown in the third row are inserted.

Wilcoxon Signed-Ranks Procedure

Step 4: Find the sum of the **ranks that are positive**. Also find the **absolute value of the sum of the negative ranks**.

EXAMPLE: The bottom row of Table 13-4 lists the signed ranks. The sum of the positive ranks is **5 + 4 + 6 + 2 + 8 + 1 + 3 = 29**. The **sum of the negative ranks is -7**, and the **absolute value of this sum is 7**. The **two rank sums are 29 and 7**.

Step 5: Let ***T* be the *smaller* of the two sums found** in Step 4. Either sum could be used, but for a simplified procedure we arbitrarily **select the smaller of the two sums**.

EXAMPLE: The data in Table 1 result in the rank sums of 29 and 7, so the smaller of those **two sums is 7**.

Step 6: Let ***n*** be the **number of pairs of data for which the difference *d* is not 0**.

EXAMPLE: The data in Table 1 have 8 differences that are not 0, so $n = 8$.

Step 7: Determine the test statistic and critical values based on the sample size, as shown in previously.

EXAMPLE: For the data in Table 1 the test statistic is $T = 7$. The **sample size is $n = 8$**, so the critical value is found in **Table A-8**. Using **a 0.05 significance level with a two-tailed test**, the critical value from Table A-8 is **4**.

TABLE A-8 Critical Values of *T* for the Wilcoxon Signed-Ranks 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)
5	*	*	*	1
6	*	*	1	2
7	*	0	2	4
8	0	2	4	6
9	2	3	6	8
10	3	5	8	11
11	5	7	11	14
12	7	10	14	17
13	10	13	17	21
14	13	16	21	26
15	16	20	25	30
16	19	24	30	36
17	23	28	35	41
18	28	33	40	47
19	32	38	46	54
20	37	43	52	60
21	43	49	59	68
22	49	56	66	75
23	55	62	73	83
24	61	69	81	92
25	68	77	90	101
26	76	85	98	110
27	84	93	107	120
28	92	102	117	130
29	100	111	127	141
30	109	120	137	152

Table A-8

NOTES:

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

2. Conclusions:

Reject the null hypothesis if the **test statistic *T* is less than or equal to the critical value** found in this table.

Fail to reject the null hypothesis if the test statistic *T* is greater than the critical value found in the table.

Step 8: Reject the null hypothesis if the sample data lead to a test statistic that is in the critical region—that is, the test statistic is less than or equal to the critical value(s). Otherwise, fail to reject the null hypothesis.

EXAMPLE: For the sample of matched pairs in the first two rows of Table 1, the **test statistic is *T* = 7** and the critical value is 4, so the test statistic is *not* less than or equal to the critical value. Consequently, we fail to reject the null hypothesis that the matched pairs are from a population.

Test-statistic $\leq$ Critical value

7 $\leq$ 4 (False)

Conclusion Table A-8 includes a note stating that we should reject the null hypothesis if the test statistic *T* is less than or equal to the critical value. Because the test statistic of ***T* = 7** is *not* less than or equal to the critical value of 4, **we fail to reject the null hypothesis.**

INTERPRETATION

We conclude that there is not sufficient evidence to support the claim that for males, there is a **significant difference between measured weights and reported weights**.

Based on the very small sample, **it appears that males do not tend to report weights that are much different from their actual weights**. It is possible that a much larger sample would lead to a different conclusion.

Data Set 5: Body Temperatures

Body temperatures (°F) 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.

TI-83/84 Plus list names (BODYTEMP): D1T8, D1T12, D2T8, D2T12 (no list for SEX and SMOKE). **Missing data values are represented by 9999.**

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 5 "Body Temperatures" includes measured body temperatures of adults. Use the 106 temperatures listed for 12 AM on Day 2 with the Wilcoxon signed-ranks test to test the claim that the median is less than 98.6^0**F**. Use a 0.05 significance level.

SOLUTION

REQUIREMENT CHECK

(1) The design of the experiment that led to the data in Data Set justifies treating the sample as a simple random sample.

(2) The requirement of an approximately symmetric distribution of differences is satisfied, because a histogram of those differences is approximately symmetric.

- By pairing each individual sample value with the median of 98.6^0**F**, we are working with matched pairs. Using the technology the test statistic of ***T* = 661**, which converts to the **test statistic *z* = -5.67**. (The display is from a two-tailed test; for this left-tailed test, the critical value is -1.645.)

- The test statistic of z = -5.67 yields a P-value of 0.000, so we reject the null hypothesis that the population of differences between body temperatures and the claimed median of 98.6^0**F** is zero.

- There is sufficient evidence to support the claim that the median body temperature is less than 98.6^0**F**. This is the same conclusion that results from the sign test.

Wilcoxon rank-sum test

The **Wilcoxon rank-sum test** is a nonparametric test that uses ranks of sample data from **two independent populations** to test this null hypothesis:

***H_0*: ** Two independent samples come from populations with equal medians.

(The **alternative hypothesis *H_1*** can be any one of the following three possibilities:

The two populations have *different* medians, or **the first population has a median *greater than* the median of the second population**, or **the first population has a median *less than* the median of the second population**.)

Wilcoxon Rank-Sum Test

Objective

Use the Wilcoxon rank-sum test **with samples from two independent populations** for the following null and alternative hypotheses:

***H_0*:** The two samples come from populations with equal medians.

***H_1*:** The median of the first population is different from (or greater than, or less than) the median from the second population.

Notation

n_1 = size of Sample 1

n_2 = size of Sample 2

R_1 = sum of ranks for Sample 1

R_2 = sum of ranks for Sample 2

R = same as R_1 (sum of ranks for Sample 1)

μ_R = mean of the sample R values that is expected **when the two populations have equal medians**

σ_R = standard deviation of the sample R values that is expected **with two populations having equal medians**

Requirements

1. There are two independent simple random samples.

2. Each of the two samples has **more than 10 values**.

(For samples with 10 or fewer values, special tables are available in special reference books, such as *CRC Standard Probability and Statistics Tables and Formulae,* published by CRC Press.)

Note: There is *no* requirement that the two populations have a normal distribution or any other particular distribution.

Test Statistic

$$z = \frac{R - \mu_R}{\sigma_R}$$

Where

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} \quad \text{and} \quad \sigma_R = \sqrt{\frac{n_1\, n_2(n_1 + n_2 + 1)}{12}}$$

n_1 = size of the sample from which the rank sum R is found

n_2 = size of the other sample

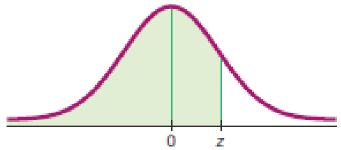
R = sum of ranks of the sample with size n_1

P-Values

P-values can be found from technology or by using the z test statistic and Table A-2.

Critical Values

Critical values can be found in Table A-2 (because the test statistic is based on the normal distribution).



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

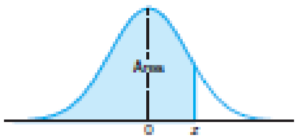


Table A.3 Areas under the Normal Curve

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

TABLE 13-5 Heights (mm) of Males from ANSUR I and ANSUR II

ANSUR I 1988	ANSUR II 2012
1698 **(5)**	1810 **(21)**
1727 **(8)**	1850 **(25)**
1734 **(11)**	1777 **(16)**
1684 **(3)**	1811 **(22)**
1667 **(1)**	1780 **(17.5)**
1680 **(2)**	1733 **(10)**
1785 **(19)**	1814 **(23)**
1885 **(27)**	1861 **(26)**
1841 **(24)**	1709 **(7)**
1702 **(6)**	1740 **(13)**
1738 **(12)**	1694 **(4)**
1732 **(9)**	1766 **(15)**
	1748 **(14)**
	1794 **(20)**
	1780 **(17.5)**
$n_1 = 12$	$n_2 = 15$
$R_1 = 127$	$R_2 = 251$

Procedure for Finding the Value of the Test Statistic

To see how the following steps are applied, refer to the sample data listed in Table 13-5 on the next slides. The data are heights (mm) of males randomly selected from Data Set 2 "ANSUR I 1988" and Data Set 3 "ANSUR II 2012."

Step 1: Temporarily combine the two samples into one big sample, then replace each **sample value with its rank**. (The lowest value gets a rank of 1, the next lowest value gets a rank of 2, and so on. If values are tied, assign to them the mean of the ranks involved in the tie.

EXAMPLE: In Table 13-5, the ranks of the 27 male heights are shown in parentheses. The **rank of 1** is assigned to **the lowest sample value of 1667**, the **rank of 2** is assigned to the next **lowest value of 1680**, and the **rank of 3** is assigned to the next **lowest value of 1684**. The **17th and 18th values are tied at 1780**, so we assign **the rank of 17.5 to both of those tied values.**

Step 2: Find the sum of the ranks for either one of the two samples.

EXAMPLE: In Table 13-5, the sum of the ranks from the first **sample is 127**. (That is, R_1 = 5 + 8 + 11 + g + 9 = 127.) The sum of the ranks from the second sample is **251**

Step 3: Calculate the value of the *z* test statistic as shown in the preceding slides, where either sample can be used as “Sample 1.” (If both sample sizes are greater than 10, then the sampling distribution of ***R* is approximately normal with mean μ_R and standard deviation σ_R**, and the test statistic is as shown in the preceding previous slides.)

TABLE 13-5 Heights (mm) of Males from ANSUR I and ANSUR II

ANSUR I 1988	ANSUR II 2012
1698 **(5)**	1810 **(21)**
1727 **(8)**	1850 **(25)**
1734 **(11)**	1777 **(16)**
1684 **(3)**	1811 **(22)**
1667 **(1)**	1780 **(17.5)**
1680 **(2)**	1733 **(10)**
1785 **(19)**	1814 **(23)**
1885 **(27)**	1861 **(26)**
1841 **(24)**	1709 **(7)**
1702 **(6)**	1740 **(13)**
1738 **(12)**	1694 **(4)**
1732 **(9)**	1766 **(15)**
	1748 **(14)**
	1794 **(20)**
	1780 **(17.5)**
$n_1 = 12$	$n_2 = 15$
$R_1 = 127$	$R_2 = 251$

Data Set 2: ANSUR I 1988

ANSUR is an abbreviation of "anthropometric survey." The ANSUR I study was conducted in 1988. (See also the following ANSUR II data set.) This data set consists of body measurements from 3982 U.S. Army personnel (first five rows shown here, not all data columns shown). **AGE** is in years, **WEIGHT** is in kilograms (kg), for **GENDER** 1 = male and 0 = female, for **WRITING HAND** 1 = right and 2 = left and 3 = both, and the other body measurements are in millimeters (mm). Additional detail on body measurements in this data set can be found at TriolaStats.com/ansur. Data are from the U.S. Army.

TI-83/84 list names (ANSUR1)*: A1AGE, A1HT, A1WGT, A1FTL, A1HC, A1CC, A1NC, A1WC, A1SW, A1SH, A1SKH, A1SEH, A1NHT, A1PPD, A1ARM, A1WH, A1GND

**NOTE: TI lists are limited to 500 rows due to calculator memory constraints.*

AGE	HEIGHT	WEIGHT	FOOT LENGTH	HEAD CIRC	CHEST CIRC	SHOULDER WIDTH	SITTING HT	ARM SPAN	WRITING HAND	GENDER (1 = M)
34	1735	88.3	260	572	1052	490	888	1813	1	1
37	1830	86.5	290	590	1029	485	905	1916	1	1
38	1726	71.3	254	572	995	500	907	1827	1	1
33	1783	81.6	271	593	966	484	948	1846	1	1
42	1669	75.6	240	546	1032	479	856	1712	1	1

Data Set 3: ANSUR II 2012

ANSUR is an abbreviation of "anthropometric survey." The ANSUR II study was conducted in 2012. (See also the preceding ANSUR I data set.) This data set consists of body measurements from 6068 U.S. Army personnel (first five rows shown here, not all data columns shown). **AGE** is in years, **WEIGHT** is in kilograms (kg), for **GENDER** 1 = male and 0 = female, for **WRITING HAND** 1 = right and 2 = left and 3 = both, and the other body measurements are in millimeters (mm). Additional detail on body measurements can be found at TriolaStats.com/ansur. Data are from the U.S. Army.

TI-83/84 list names (ANSUR2)*: A2AGE, A2HT, A2WGT, A2FTL, A2HC, A2CC, A2NC, A2WC, A2SW, A2SH, A2SKH, A2SEH, A2NHT, A2PPD, A2ARM, A2WH, A2GND

**NOTE: TI lists are limited to 500 rows due to calculator memory constraints.*

AGE	HEIGHT	WEIGHT	FOOT LENGTH	HEAD CIRC	CHEST CIRC	SHOULDER WIDTH	SITTING HT	ARM SPAN	WRITING HAND	GENDER (1 = M)
41	1776	81.5	273	583	1074	493	928	1782	1	1
35	1702	72.6	263	568	1021	479	884	1745	2	1
42	1735	92.9	270	573	1120	544	917	1867	2	1
31	1655	79.4	267	576	1114	518	903	1708	1	1
21	1914	94.6	305	566	1048	524	919	2035	1	1

Example: Heights of Males from ANSUR I 1988 and ANSUR II 2012

Table 13-5 lists samples of heights of males from the ANSUR I 1988 and ANSUR II 2012 data sets. Use a 0.05 significance level to test the claim that the two samples are **from populations with the same median**.

SOLUTION

REQUIREMENT CHECK

(1) The sample data are **two independent simple random samples**.

(2) The **sample sizes are 12 and 15**, so both **sample sizes are greater than 10.**

The requirements are satisfied.

The null and alternative hypotheses are as follows:

***H_0*:** The two samples are from populations with the same median.

***H_1*:** The two samples are from populations with different medians.

Rank the combined list of all 27 male heights, beginning with a rank of 1 (assigned to the lowest value of 1667). The ranks corresponding to the individual sample values are shown in parentheses in Table 13-5. R denotes the sum of the ranks for the sample we choose as Sample 1.

If we choose the ANSUR I 1988 sample, we get R = 5 + 8 + 11 + g + 9 = 127

Because there are 12 heights in the first sample, we have n_1 = 12. Also, n_2 = 15 because there are 15 heights in the second sample

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$= \frac{12(12 + 15 + 1)}{2}$$

$$=168$$

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$$= \sqrt{\frac{(12)(15)(12 + 15 + 1)}{12}}$$

$$=20.4939$$

$$z = \frac{R - \mu_R}{\sigma_R}$$

$$= \frac{127 - 168}{20.4939}$$

$$= -2.00$$

The test is two-tailed because a **large positive value of *z*** would indicate that disproportionately more higher ranks are found in Sample 1, and a large negative value of z would indicate that disproportionately more lower ranks are found in Sample 1. In either case, we would have strong evidence against the claim that the two samples come from populations with equal medians.

The significance of the test statistic z can be treated as in previous chapters. We are testing (with a = 0.05) the hypothesis that the **two populations have equal medians**, so we have a **two-tailed test**.

***P*-Value:** Using the unrounded z score, the P-value is 0.045, so we reject the null hypothesis that the two samples are from populations with the same median.

Critical Values: If we use the critical values of $\boldsymbol{z = \pm 1.96}$, we see that the **test statistic of $\boldsymbol{z}$ = -2.00** does fall within the critical region, so we reject the null hypothesis that the **two samples are from populations with the same median**

INTERPRETATION

- There is sufficient evidence to warrant rejection of the claim that the sample of male heights from ANSUR I 1988 and the sample of male heights from ANSUR II 2012 are from populations with the same median. **It appears that the medians are different**.
- Based on the listed sample data, **it appears that the heights from 1988 are different than the heights from 2012**. Because the **heights from 2012 have a larger median**, it appears that **males became taller as time passed from 1988 to 2012**, although a one-sided hypothesis test would be better for testing that conclusion. A one-sided test would lead to the conclusion that the 2012 heights are significantly larger than the 1988 heights.