NON-PARAMETRIC STATISTICAL TESTS

rev Feb 2016

Scientific research is centered around the formulation of hypotheses. Statistical tests determine if the results you obtain differ from what would be obtained by chance alone. The *non-parametric tests* given below can be used when sample sizes are small or you are unsure if the data are normally distributed.

MANN-WHITNEY U TEST -- *compare two sample medians*

Purpose: To test if the medians of two treatments are equal; does not assume data is normally distributed, or that variances are equal. If data is normal (i.e., large sample size) use a t-test (parametric).

Procedure:

1) List all observations in two columns, each corresponding to a treatment
2) Rank all observations combined from smallest to largest.
3) In case of ties, assign the average rank to tied observations, then skip to next unused rank
4) Sum the *ranks* for each column of treatment observations (R_i)
5) Calculate the U-statistic for each sample.

$U_1 = (N_1)(N_2) + [N_1(N_1 + 1)/2] - R_1$

$U_2 = (N_1)(N_2) + [N_2(N_2 + 1)/2] - R_2$ *check: [$U_1 + U_2$] should = [$N_1 * N_2$]*

6) Compare the *smaller* of the $U_{calculated}$ with the $U_{critical}$ from the U-table of critical values.

7) If $U_{calc} <= U_{crit}$, reject H_o *note: some U_{crit} tables require **larger** U_{calc} **>** U_{crit}*

Sample data set: ***plus calcs for Wilcoxon***

Compare the median number of bromeliads counted on 7 trees with smooth bark and 8 with rough bark.

Tree #	Smooth Bark	*rank1*	Rough Bark	*rank2*
1	2	*1*	12	*12*
2	5	*3*	14	*14*
3	12	*12*	8	*6.5*
4	8	*6.5*	7	*4.5*
5	7	*4.5*	12	*12*
6	4	*2*	11	*10*
7	9	*8*	17	*15*
8			10	*9*
median	***7***		***11.5***	
mean	*6.71*		*11.38*	
rank sum		***37.0***		***83.0***

Diff	*rankD*	
10	*7*	
9	*6*	
-4	*2*	
-1	*1*	
5	*3*	
7	*4*	
8	*5*	
x		
	3	***T-***
	25	***T+***

U1 = (8)(7) + [7(8)/2] - 37 = 56 + 28 - 37.0 = **47.0**

U2 = (8)(7) + [8(9)/2] - 83 = 56 + 36 - 83.0 = **9.0** ***{*check: 47.0+9.0 = 7*8 = 56}***

Smallest is U_1 = **9.0 = Ucalc**

find Ucrit: check U table (p = 0.05), with N1 = 7, N2 = 8 ... **Ucrit = 10**

Ucalc < Ucrit, so reject H_0 -- there *is* a significant difference

"On average, fewer bromeliads were found on smooth-barked trees (median 7 per tree, n=7) than on rough-barked trees (median 11.5 per tree, n=8), a significant difference at p < 0.05 (Mann-Whitney Ucalc=9.0)."

WILCOXON SIGN-RANK TEST -- *compare medians of two* paired *samples*

Purpose: To test if the medians of two treatments are equal, when data were collected from paired samples: e.g., count epiphytes 0-1 m from the ground on a tree vs. 1-2 m from the ground *on same tree*. *Note: Must have 5 or more pairs. If data are normal, use paired t-test

Very powerful design, since variability of many other (non-measured) factors is controlled. But therefore must be careful that experimental design actually *is* paired before using Wilcoxon.

Procedure:

1) List all paired data (*n* pairs) in two columns, one for each treatment
2) Calculate a column of the difference between each measurement pair; keep pluses and minuses
3) Ignore all zeros, ignore sign (+/-), and rank the absolute value of each difference
4) Sum the *ranks* for all differences with a minus ($\mathbf{T_-}$), and sum all ranks for positive differences ($\mathbf{T_+}$)
5) Select the *smaller* of $\mathbf{T_-}$ and $\mathbf{T_+}$ as your calculated statistic ($\mathbf{T_{calc}}$)
   *note: if all differences are unidirectional (i.e., no negatives), then $T_{calc} = 0$

6) If $T_{calc} <= T_{crit}$, reject H_0 (at that highest level of significance for T_{crit})

*see calculations on previous page: $\mathbf{T_{calc} = 3}$; $\mathbf{T_{crit} = 3}$ (n=7, p=0.05); **reject H_0**

KRUSKAL-WALLIS TEST -- *compare K treatment medians*

Purpose: To test if the medians of three or more treatments are equal; does not assume normality or equal variance between treatments. If distribution is normal, use one-way Analysis of Variance (parametric).

Procedure:

1) List observations for the K treatments, with $n_i \ldots n_k$ observations in each.
2) Rank all the combined observations from smallest to largest, as described for Mann-Whitney U)
3) Sum the ranks in each treatment, R_i to R_k.
4) Calculate the Kruskall-Wallis statistic (H) as follows:

$$H = \frac{12}{N(N+1)} * \left[\sum\left(R_i^2 / n_i\right)\right] - 3(N+1)$$

Where K = # of groups or treatments
n_i = the number of observations in the i^{th} treatment
N = the total number of observations (over all treatments)
R_i = the sum of the ranks in the i^{th} treatment

5) Compare to χ^2 critical value table with v = K - 1

6) If H_{calc} is > χ^2_{crit}, reject H_o

CORRELATION: *testing the* significance *of the coefficient of correlation*

The following test converts an $\mathbf{r^2}$ (coefficient of correlation, a measure of how well a line "fits" the data points) into a t-test, thus providing a level of significance (p-value) of the confidence of the linear correlation. Obviously, the more points (**n**) the stronger the confidence. *Note, "standard" (i.e., Pearson's) correlation is *very* robust, and the non-parametric technique (Spearman Rank) is no longer in use.

$$t_{calc} = \frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

where ***r*** = square root of $\mathbf{r^2}$ = square root of Rs

Compare $\mathbf{t_{calc}}$ to $\mathbf{t_{crit}}$ (in a table of critical t-test values) with degrees of freedom (df) = n-2. If $\mathbf{t_{calc}}$ exceeds the critical value, then the two variables are significantly correlated.

CHI-SQUARE (χ^2) TEST -- *compare* observed *against* expected *frequencies*

Use when both dependent & independent variables are categorical (i.e., species and substrate); data are *counts of frequencies*. Not appropriate for nominal data. Assumes samples are random and independent.

Contingency Test: To test for the independence of two factors, A and B; said another way, it tests for contingency (aka dependency) between factors. Null hypothesis (H_O) is that factors A and B are *independent*.

Goodness-of-fit Test: To test if observed frequencies (counts) differ from expectations under null hypothesis (e.g., we expect a tossed coin to show heads 50%; if actual data is 65%, is that non-random?).

Procedure (Contingency):

1) Arrange *Observed* data as below with each cell containing the frequencies (# of occurrences) of each observation. Add up all subsamples (traps, plots within sites, etc) to calculate frequencies.

Fungus Color	**Fungus Beetle Color**		
	black	white	*totals*
white	**22**	**38**	*60*
black	**31**	**18**	*49*
totals	*53*	*56*	***109***

2) Calculate the sums for each row and column, and the **grand total** (add all row totals)
3) Calculate the *Expected* frequencies (*E*) for each cell as follows:
   E of A,1 = (Column sum A) x (Row sum 1) / TOTAL

   Thus *E* of upper-left cell (Black Beetle - White Fungus) is 53 x 60 / 109 = **29.2**

Expected frequencies:

	black beetle	white beetle
white fungus	**29.2**	**30.8**
black fungus	**23.8**	**25.2**

4) Compare the **O**bserved versus **E**xpected values to determine the χ^2 statistic:
   $\chi^2 = \Sigma [(\mathbf{O - E})^2 \ / \mathbf{E}] = 1.77 + 1.68 + 2.18 + 2.06 =$ **7.69**

   {degrees of freedom, *v* = (# of rows - 1)(# of columns - 1) = 1}

5) Compare the χ^2 calculated to the χ^2 critical value at the appropriate degrees of freedom:

6) If $\chi^2_{calc} > \chi^2_{crit}$, reject H_O (at that highest level of significance for χ^2_{crit})

   $\chi^2_{calc} = 7.64 > \chi^2_{crit,\ v=1} = 6.64$ at $\alpha = 0.01$ (*α = p = the probability that the result is due to chance*)

7) The test rejects the null hypothesis. You would write this result as "there was a significant contingency between beetle color and fungus color (χ^2_{calc}=7.64, v=1, p < 0.01)."

Procedure (Goodness-of-Fit): *when χ^2 has only one row (usually)*

1) Arrange *Observed* data with each cell containing the frequencies (# of occurrences).
   example: do hummingbirds visit red, yellow, white flowers non-randomly? Watch 10 of each flower.
2) Formulate the null hypothesis: often, H_0 states "counts will be randomly distributed"
3) Calculate the *Expected* values based on null hypothesis:
   example: total hummingbird visits should be evenly divided among 3 flower colors (33% each)
4) Calculate χ^2 statistic as above (using Observed and Expected), compare to χ^2 at v=1.

Critical Values for the Mann-Whitney U-Test

Level of significance: 5% (P = 0.05)

Size of the largest sample (n_2)

Size of the smallest sample (n_1)

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11	12	13	13
4	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14	15	16	17	17	18	19	20	21	22	23
5	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20	22	23	24	25	27	28	29	30	32	33
6		5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	29	30	32	33	35	37	38	40	42	43
7			8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54
8				13	15	17	19	22	24	26	29	31	34	36	38	41	43	45	48	50	53	55	57	60	62	65
9					17	20	23	26	28	31	34	37	39	42	45	48	50	53	56	59	62	64	67	70	73	76
10						23	26	29	33	36	39	42	45	48	52	55	58	61	64	67	71	74	77	80	83	87
11							30	33	37	40	44	47	51	55	58	62	65	69	73	76	80	83	87	90	94	98
12								37	41	45	49	53	57	61	65	69	73	77	81	85	89	93	97	101	105	109
13									45	50	54	59	63	67	72	76	80	85	89	94	98	102	107	111	116	120
14										55	59	64	67	74	78	83	88	93	98	102	107	112	118	122	127	131
15											64	70	75	80	85	90	96	101	106	111	117	122	125	132	138	143
16												75	81	86	92	98	103	109	115	120	126	132	138	143	149	154
17													87	93	99	105	111	117	123	129	135	141	147	154	160	166
18														99	106	112	119	125	132	138	145	151	158	164	171	177
19															113	119	126	133	140	147	154	161	168	175	182	189
20																127	134	141	149	156	163	171	178	186	193	200
21																	142	150	157	165	173	181	188	196	204	212
22																		158	166	174	182	191	199	207	215	223
23																			175	183	192	200	209	218	226	235
24																				192	201	210	219	228	238	247
25																					211	220	230	239	249	258
26																						230	240	250	260	270
27																							250	261	271	282
28																								272	282	293
29																									294	305
30																										317

The Open Door Web Site

<u>Critical Values of *T* for Wilcoxon Sign-Rank Test</u>

n	**0.10**	**0.05**	**0.02**	**0.01**
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	32
20	60	52	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68

χ^2 critical value table for Kruskall Wallis & Chi-Square tests

Upper critical values of chi-square distribution with *v* degrees of freedom
Probability of exceeding the critical value

v	0.10	0.05	0.025	0.01	0.001
1	2.706	3.841	5.024	6.635	10.828
2	4.605	5.991	7.378	9.210	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588
11	17.275	19.675	21.920	24.725	31.264
12	18.549	21.026	23.337	26.217	32.910
13	19.812	22.362	24.736	27.688	34.528
14	21.064	23.685	26.119	29.141	36.123
15	22.307	24.996	27.488	30.578	37.697
16	23.542	26.296	28.845	32.000	39.252
17	24.769	27.587	30.191	33.409	40.790
18	25.989	28.869	31.526	34.805	42.312
19	27.204	30.144	32.852	36.191	43.820
20	28.412	31.410	34.170	37.566	45.315

Lower critical values of chi-square distribution with *v* degrees of freedom

Probability of exceeding the critical value

v	0.90	0.95	0.975	0.99	0.999
1.	.016	.004	.001	.000	.000
2.	.211	.103	.051	.020	.002
3.	.584	.352	.216	.115	.024
4.	1.064	.711	.484	.297	.091
5.	1.610	1.145	.831	.554	.210
6.	2.204	1.635	1.237	.872	.381
7.	2.833	2.167	1.690	1.239	.598
8.	3.490	2.733	2.180	1.646	.857
9.	4.168	3.325	2.700	2.088	1.152
10.	4.865	3.940	3.247	2.558	1.479
11.	5.578	4.575	3.816	3.053	1.834
12.	6.304	5.226	4.404	3.571	2.214
13.	7.042	5.892	5.009	4.107	2.617
14.	7.790	6.571	5.629	4.660	3.041
15.	8.547	7.261	6.262	5.229	3.483
16.	9.312	7.962	6.908	5.812	3.942
17.	10.085	8.672	7.564	6.408	4.416
18.	10.865	9.390	8.231	7.015	4.905
19.	11.651	10.117	8.907	7.633	5.407
20.	12.443	10.851	9.591	8.260	5.921

Upper critical values of Student's t distribution

(Probability of exceeding the critical value)

d.f	0.10	0.05	0.025	0.01	0.005	0.001
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499
9.	1.383	1.833	2.262	2.821	3.250	4.296
10.	1.372	1.812	2.228	2.764	3.169	4.143
11.	1.363	1.796	2.201	2.718	3.106	4.024
12.	1.356	1.782	2.179	2.681	3.055	3.929
13.	1.350	1.771	2.160	2.650	3.012	3.852
14.	1.345	1.761	2.145	2.624	2.977	3.787
15.	1.341	1.753	2.131	2.602	2.947	3.733
16.	1.337	1.746	2.120	2.583	2.921	3.686
17.	1.333	1.740	2.110	2.567	2.898	3.646
18.	1.330	1.734	2.101	2.552	2.878	3.610
19.	1.328	1.729	2.093	2.539	2.861	3.579
20.	1.325	1.725	2.086	2.528	2.845	3.552
21.	1.323	1.721	2.080	2.518	2.831	3.527
22.	1.321	1.717	2.074	2.508	2.819	3.505
23.	1.319	1.714	2.069	2.500	2.807	3.485
24.	1.318	1.711	2.064	2.492	2.797	3.467
25.	1.316	1.708	2.060	2.485	2.787	3.450
26.	1.315	1.706	2.056	2.479	2.779	3.435
27.	1.314	1.703	2.052	2.473	2.771	3.421
28.	1.313	1.701	2.048	2.467	2.763	3.408
29.	1.311	1.699	2.045	2.462	2.756	3.396
30.	1.310	1.697	2.042	2.457	2.750	3.385
31.	1.309	1.696	2.040	2.453	2.744	3.375
32.	1.309	1.694	2.037	2.449	2.738	3.365
33.	1.308	1.692	2.035	2.445	2.733	3.356
34.	1.307	1.691	2.032	2.441	2.728	3.348
35.	1.306	1.690	2.030	2.438	2.724	3.340
36.	1.306	1.688	2.028	2.434	2.719	3.333
37.	1.305	1.687	2.026	2.431	2.715	3.326
38.	1.304	1.686	2.024	2.429	2.712	3.319
39.	1.304	1.685	2.023	2.426	2.708	3.313
40.	1.303	1.684	2.021	2.423	2.704	3.307
41.	1.303	1.683	2.020	2.421	2.701	3.301
42.	1.302	1.682	2.018	2.418	2.698	3.296
43.	1.302	1.681	2.017	2.416	2.695	3.291
44.	1.301	1.680	2.015	2.414	2.692	3.286
45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265
50.	1.299	1.676	2.009	2.403	2.678	3.261
51.	1.298	1.675	2.008	2.402	2.676	3.258
52.	1.298	1.675	2.007	2.400	2.674	3.255
53.	1.298	1.674	2.006	2.399	2.672	3.251
54.	1.297	1.674	2.005	2.397	2.670	3.248
55.	1.297	1.673	2.004	2.396	2.668	3.245