Lecture 14

Hypothesis Testing (Tests of Significance) – 2 Sample Hypothesis Tests

So far, we have only been considering hypothesis tests of 1 sample. This means that we look at 1 sample in isolation and make inference about the corresponding population parameters (μ, p).

Visually:

Now, we are going to consider cases where we have 2 samples from 2 different populations and wish to compare the parameters of these populations.

Visually:

The key difference is:

Comparing Two Population Means

Before laying out the groundwork for a 2-sample comparison of means as we did with the 3 different 1-sample tests, we must first make an important distinction between 2 cases:

Independent Samples →

Matched Pairs →

Example – Identifying Independence

For the following examples identify if we are discussing a 1-sample situation, or 2-sample situation. If we are discussing a 2-sample situation, identify whether the samples are matched pairs or independent samples.

a. Researchers are interested in testing the effect of drinking alcohol on driving. So, they measure the time taken for 30 subjects each to complete a driving course. Then, each subject drinks 2 beers in 5 minutes and the researcher times the driving course again.

b. Researchers are interested in testing the effect of drinking alcohol on driving. So, they randomly select 30 U.S. males to complete a timed driving course. Then, they randomly select another 30 males to each drink 2 beers in 5 minutes and take the driving course.

c. Researchers are interested in testing the effect of drinking alcohol on driving. So, they randomly select 30 females and measure the time taken for each to complete a driving course. The goal of the study is to see if female times are lower than the national average of 1.34 minutes.

The Independent Samples Case – Two-Sample T-Test

As always, we lay out the 4 fundamental steps in hypothesis testing:

Safe to use if:

i.

ii.

iii.

1. Hypothesis:

   H_0:

   H_A:

Note: You MUST label each mu!!!

2. Test statistic:

   $t_0 =$

3. P-value →

   If $H_A: \mu_1 < \mu_2$ →

   If $H_A: \mu_1 > \mu_2$ →

   If $H_A: \mu_1 \neq \mu_2$ →

4. Conclusion (in context)

Example – Childhood Weight

There is concern that childhood obesity is increasing with time. Is there really cause for concern (or is it just hype)? To help provide data to answer such questions, the U.S. government sponsors the National Health and Nutrition Examination Survey (NHANES), a periodically-updated record of many health variables of interest. According to a recent NHANES, the weight (in pounds) of a sample of 10-year-old males from recent years, compared to a sample of 10-year-old males in the 1960's, is summarized as follows:

	n	$\bar{x}$	s
10-year-old males Recent Years	187	84.9	24.6
10-year-old males 1963-1965	576	74.2	14.4

Analyze these data to formally test whether 10-year-old males from recent years weigh more than in the 1960's at an alpha significance level of .05 by answering the following steps.

a. What are the hypotheses of interest?

b. Why can we use the two-sample T-test here?

c. Calculate the test statistic (including the degrees of freedom).

d. Calculate the p-value, and interpret it in the context of this question.

e. Draw your conclusions in context at the alpha significance level of .05.

Again, we can use a confidence interval approach for the two-sided alternative:

To find the appropriate t-score, go to the t-score table and find the degrees of freedom you have, as well as the desired probability of $\frac{1-CC}{2} + CC$, and take that t-score for your formula.

Note: Minitab only provides the full confidence interval as part of the output for the two-sided test.

Interpreting the Confidence Interval:

We are ___% confident that the mean_____ in <u>(sample 1)</u> is between <u>(lower bound)</u> and <u>(upper bound)</u> <u>higher/lower</u> than the mean ______ in <u>(sample 2)</u>.

Note: If the C.I. contains the value 0, then we will not reject H_0, and thus we cannot conclude H_A. If the C.I. does not contain the value 0, then we can reject the null hypothesis and conclude the alternative.

Now, let's look at the case where we have matched pairs…

The Matched Pairs Case – Paired T-Test

Visually:

Note: These two samples have to be the same size, otherwise you can't take their differences!

In this case, the inference is not based on two samples but rather on the ONE sample of differences $d_1, d_2, \dots, d_n$.

More specifically, the inference is based on the average of the differences:

$$\bar{x}_d = \frac{\sum_{i=1}^{n} d_i}{n}$$

and the standard deviations of the differences s_d.

If $\bar{x}_d = \frac{\sum_{i=1}^{n} d_i}{n}$ is unusually high or unusually small (as indicated by the p-value), then the data provide evidence to reject H_0.

Let's lay out formally the 4 steps for the paired t-test:

Safe to use if:

i.

ii.

iii.

1. Hypothesis:

   H_0:

   H_A:

Note: You MUST label what μ_d is (by labeling what each mu represents)!!!

2. Test statistic:

   $t_0 =$

3. P-value →

   If $H_A: \mu_d < 0$ →

   If $H_A: \mu_d > 0$ →

   If $H_A: \mu_d \neq 0$ →

4. Conclusion (in context)

Note: This is essentially just a 1-sample t-test where the 1 sample is the distribution of the differences.

Example – Susan Farber's Intelligence "Nurture" vs "Nature" Study

Researchers have long been interested in the extent to which intelligence, as measured by IQ score, is affected by "nurture" as opposed to "nature": that is, are people's IQ scores mainly a result of their upbringing environment, or are they mainly an inherited trait? A study was designed to measure the effect of home environment on intelligence, or more specifically, the study was designed to address the question: "Are there significant differences in IQ scores between people who were raised by their birth parents, and those who were raised by someone else?"

In order to be able to answer this question, the researchers needed to get two groups of subjects (one from the population of people who were raised by their birth parents, and one from the population of people who were raised by someone else) who are as similar as possible in all other respects. In particular, since genetic differences may also affect intelligence, the researchers wanted to control for this confounding factor.

We know from our discussion on study design (in the Producing Data unit of the course) that one way to (at least theoretically) control for all confounding factors is randomization – randomizing subjects to the different treatment groups. In this case, however, this is not possible. This is an observational study; you cannot randomize children to either be raised by their birth parents or to be raised by someone else. How else can we eliminate the genetics factor? We can conduct twin studies.

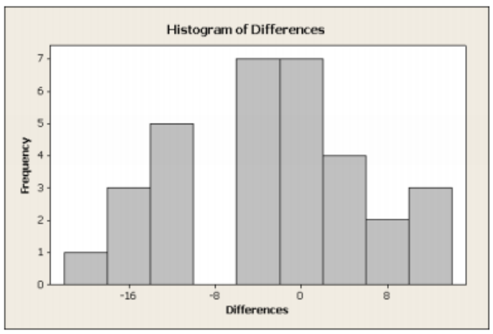
Because identical twins are genetically the same, a good design for obtaining information to answer this question would be to compare IQ scores for identical twins, one of whom is raised by birth parents and the other by someone else. Such a design (matched pairs) is an excellent way of making a comparison between individuals who only differ with respect to the explanatory variable of interest (upbringing) but are as alike as they can possibly be in all other important aspects (inborn intelligence). Identical twins raised apart were studied by Susan Farber, who published her studies in the book "Identical Twins Reared Apart" (1981, Publisher: Basic Books). In this problem we are going to use data that appear in Farber's book of 32 pairs of identical twins reared apart.

Note the following output:

Descriptive Statistics: Differences

Variable	N	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Differences	32	-2.91	1.57	8.89	-19.00	-11.00	-2.50	3.75	13.00

Pair	1	2	3	4	• • •	32
TWIN 1 (birth parents)	113	94	99	77	• • •	97
TWIN 2 (someone else)	109	100	86	80	• • •	98
Differences (twin 1 - twin 2)	4	-6	13	-3	• • •	-1



a. Identify the hypotheses of interest.

b. Explain why the paired t-test is proper here.

c. Calculate the test statistic (including the degrees of freedom).

d. Calculate the p-value.

e. What is your conclusion in context at the alpha significance level of .05?

Per usual, with the two-sided alternative we can use a confidence interval approach to the hypothesis testing:

If 0 falls inside the interval →

If 0 is not inside the interval →

The formula for the confidence interval is:

Example – Susan Farber Intelligence Study (cont.)

Use a 95% confidence interval approach to hypothesis testing in order to confirm the findings above (include the appropriate hypotheses, the interval, and your conclusion in context).

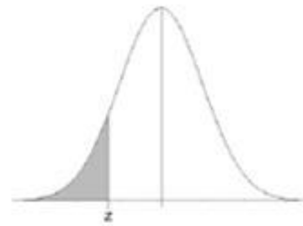
Note: If a result is not significant, all it means is that the data do not provide enough evidence to reject H_0. It does NOT mean that the result is not important, as we have seen in the group projects.

Note: We are not covering the Wilcoxon Rank Sum Test in this course, so please ignore this on Minitab.

Helpful Flowchart:

Minitab Notes:

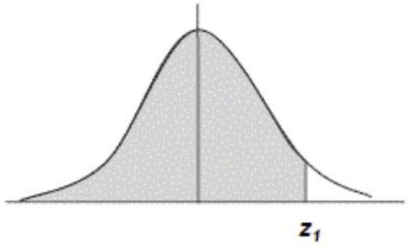
Standard Normal Cumulative Probability Table



Cumulative probabilities for NEGATIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-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

Standard Normal Distribution

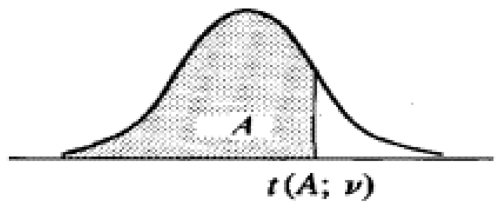


$$p(z \le z_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_1} e^{-\frac{1}{2}z^2}\, dz$$

z_1	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

T-Score Table

Entry is $t(A;\nu)$ where $P\{t(\nu) \leq t(A;\nu)\} = A$



ν	A: .60	.70	.80	.85	.90	.95	.975
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960