One and two sample test of hypothesis

1- Introduction and Definitions

An important part of inferential statistics is hypothesis testing. As with learning anything related to mathematics, it is helpful to work through several examples.
Often, the problem confronting the scientist of engineer is producing a conclusion about some scientific theories. For example, researcher may decide on the basis of experimental evidence whether coffee drinking increases the risk of cancer in humans, or a sociologist might wish to collect appropriate data to enable him or her to decide whether a person's blood type and eye color are independent variables. In addition, each must make use of experimental data and make a decision based on a data. In each case, the conjecture can be put in the form of a statistical hypothesis.

Definition: A statistical hypothesis is an assertion or conjecture concerning one or more populations.

The true or falsity of a statistical hypothesis is never known with absolute certainty unless we examine the entire population. This, of course, would be impractical in most situations. Instead, we take a random sample from the population of interest and use the data contained in this sample to provide evidence that either supports or does not support the hypothesis. Evidence from the sample which is inconsistent with the stated hypothesis leads to reject the hypothesis.

Null Hypothesis: Denoted byH_0, refers to any hypothesis we wish to test

Alternative Hypothesis: The rejection of H_0 leads to the acceptance of an alternative hypothesis, denoted by H_1

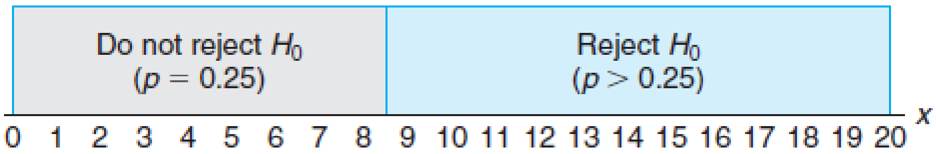
Conclusions: Reject H_0in favor of H_1 because of sufficient evidence in the data or fail to reject H_0 because of insufficient evidence in the data

Note that the conclusions do not involve a formal and literal "accept H_0".

Example: A certain type of cold vaccine is known to be only 25% effective after a period of 2 years. To determine if a new vaccine is superior in providing protection against the same virus for a longer period of time, suppose that 20 people are chosen at random and inoculated. If more than 8 of those receiving the new vaccine surpass 2-year period without contracting the virus, the new vaccine will be considered superior to the one presently in use.

2- The test statistic

The test statistic on which we base our decision is X, the number of individuals in our test group who receive protection from the new vaccine for a period of at least 2 years. The possible values possible values of X, from 0 to 20, are divided into two groups: those numbers less than or equal to 8 and those greater than 8. All possible scores greater than 8, constitute the critical region. The last number that we observe in passing into the critical region is called the **critical value**. In our illustration, the critical value is the number 8. Therefore, if $x > 8$, we reject H_0 in favor of the alternative hypothesis H_1. If $x \leq 8$, we fail to reject H_0. This decision criterion is illustrated in the figure.



The decision procedure could lead to either of two wrong conclusions. For instance, the new vaccine may be no better than the one now in use (H_0 $true$) and yet, in this particular randomly selected group of individuals, more than 8 surpass the 2-year period without contracting the virus. We would be committing an error by rejecting H_0 in favor of H_1 when, in fact, H_0 is true. Such an error is called a **type I error**.

Definition: rejection of the null hypothesis when it is true is called a **type I error.**

A second kind of error is committed if 8 or fewer of the group surpass the 2-year period successfully and we are unable to conclude that the vaccine is better when it actually is better (H_1 $true$). Thus, in this case, we fail to reject H_0 when in fact H_0 in fact. This is called **type II error**.

Definition: Non-rejection of the null hypothesis when it is false is called a **type II error**.

	H_0 is $true$	H_0 is $false$
Do not $reject$ H_0	Correct decision	Type II error
$Reject$ H_0	Type I error	Correct decision

The probability of type I error (α):

The probability of a type II error ($\boldsymbol{\beta}$):

The probability of committing a type II error, denoted by $\boldsymbol{\beta}$, is impossible to compute unless we have a specific alternative hypothesis. If we test the null hypothesis that $p = 0.25$ against the alternative hypothesis that $p = 0.5$, them we are able to compute the probability of not rejecting H_0 when it is false, in this case

Definition: the power of a test is the probability of rejecting H_0 given that a specific alternative is true. The power of a test can be computed as $1 - \beta$.

How to Avoid Errors

Type I and type II errors are part of the process of hypothesis testing. Although the errors cannot be completely eliminated, we can minimize one type of error.
Typically, when we try to decrease the probability one type of error, the probability for the other type increases.
We could decrease the value of alpha from 0.05 to 0.01, corresponding to a 99% level of confidence. However, if everything else remains the same, then the probability of a type II error will nearly always increase.
Many times the real world application of our hypothesis test will determine if we are more accepting of type I or type II errors. This will then be used when we design our statistical experiment.

3- One and two tailed tests

A test of any statistical hypothesis is called a **one-tailed test** where the alternative is one sided

, such as:

A test of any statistical hypothesis is called a **two-tailed test** where the alternative is two sided, such as:

Example: a manufacturer of a certain brand of rice cereal claims that the average saturated fat content does not exceed 1.5 grams per serving. State the null and alternative hypothesis.

Example: A real estate agent claims that 60% of all private residences being built today are 3-bedroom homes. To test this claim, a large sample of new residences is inspected; the proportion of these homes with 3-bedrooms is recorded and used as the test statistic. State the null and alternative.

4- The Use of P-Value for Decision Making in Testing Hypothesis

In testing hypothesis in which the test statistic is discrete, the critical region may be chosen arbitrarily and its size determined. Over a number of generations of statistical analysis, it had become customary to choose an α of 0.05 or 0.01 and select the critical region accordingly.

One-tailed test:

If $p-value < \alpha$ then H_0 will be rejected

If $p-value > \alpha$ then H_0 will not be rejected

Two-tailed test:

If $p-value < {}^{\alpha}/_{2}$ then H_0 will be rejected

If $p-value > {}^{\alpha}/_{2}$ then H_0 will not be rejected

Approach to Hypothesis Testing with Fixed Probability of Type I Error

- State the null and alternative hypothesis
- Choose a fixed significance level α
- Choose an appropriate test statistic and establish the critical region based on α
- Reject H_0 if the computed tests statistic is in the critical region. Otherwise, do not reject
- Draw scientific or engineering conclusions

Significance Testing (P-Value Approach)

- State null and alternative hypothesis
- Choose an appropriate test statistic
- Compute the P-value based on the computed value of the test statistic

- Use judgment based on the P-value and knowledge of the scientific system

5- Single Sample: Tests Concerning a Single Mean (Variance Known)

In this section, we formally consider tests of hypothesis on a single population mean

$$p\left(-z\alpha_{/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z\alpha_{/2}\right) = 1-\alpha$$

$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ if $-z\alpha_{/2} < z < z\alpha_{/2}$, do not reject H_0, rejection of H_0, implies acceptance of the alternative hypothesis $\mu \neq \mu_0$

Tests of one-sided hypothesis:

$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$
$H_1: \mu > \mu_0$	$H_1: \mu < \mu_0$
$If\ z > z_\alpha$	$If\ z < z_\alpha$
$If\ \alpha > P - Value, reject\ H_0$	$If\ \alpha > P - Value, reject\ H_0$

Example: A random sample of 100 recorded deaths in the United states during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

Example: A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5 kilogram. Test the hypothesis that $\mu = 8$ kilograms against the alternative that $\mu \neq 8$ kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms. Use a 0.01 level of significance.

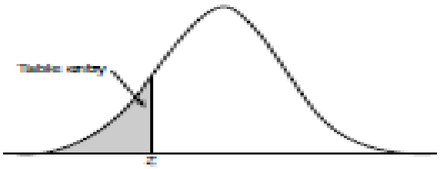


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

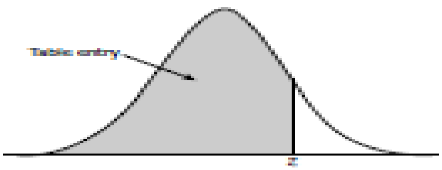


Table entry for z is the area under the standard normal curve to the left of z.

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

6- The t statistic for a test on a single mean (Variance Unknown)

For the two sided hypothesis

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

We reject H_0 at significance level of α when the computed t-statistic $t = \frac{\bar{x}-\mu.}{s/\sqrt{n}}$ exceeds $t\alpha_{/2,n-1}$ or is less than $-t\alpha_{/2,n-1}$

Example: The Edison Electric Institute has published figures on the number of kilowatt hours used annually by various home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

7- Two samples: tests on two Means with known variances

Two independent random samples of sizes n_1 and n_2, respectively, are drown from two populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$H_0: \mu_1 - \mu_2 = d_0$

$H_1: \mu_1 - \mu_2 \neq d_0$

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

8- Two samples: tests on two Means with unknown but equal variances

Two independent random samples of sizes n_1 and n_2, respectively, are drown from two populations with means μ_1 and μ_2 and sample variances s_1^2 and s_2^2.

$H_0: \mu_1 - \mu_2 = d_0$

$H_1: \mu_1 - \mu_2 \neq d_0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Example: An experiment was performed to compare the abrasive wear of two different laminated materials. 12 pieces of material 1 were tested by exposing each peace to a machine measuring wear. 10 pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 with a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units? Assume that the populations to be approximately normal with equal variances.

9- Two samples: tests on two Means with unknown but unequal variances

Two independent random samples of sizes n_1 and n_2, respectively, are drown from two populations with means μ_1 and μ_2 and sample variances s_1^2 and s_2^2.

$H_0: \mu_1 - \mu_2 = d_0$

$H_1: \mu_1 - \mu_2 \neq d_0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad v = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}}$$

The test procedure is to not reject H_0 when $-t_{\alpha/2, n-1} < t < t_{\alpha/2, n-1}$

10- Paired Observations

Testing of two means can be accomplished when data are in the form of paired observations for $\mu_1 - \mu_2$ and $t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$

Example: in a study conducted in the forestry and wildlife Department at Virginia Tech, J. A. Wesson examined the influence of the drug succinylcholine on the circulation levels of androgens in the blood. Blood samples were taken from wild, free-ranging deer immediately after they had received an intramuscular injection of succinylcholine administered using darts and a capture gun. A second blood sample was obtained from each deer 30 minutes after the first sample, after which the deer was released. The levels of androgens at time of capture and 30 minutes later, measured in Nano grams per millimeter, for 15 deer are given in the table. Assuming that the population of androgen levels at time of injection and 30 minutes later are normally distributed, test at the 0.05 level of significance whether the androgen concentrations are altered after 30 minutes.

Androgen		
Deer	At time of injection	30 minutes after injection
1	2.76	7.02
2	5.18	3.1
3	2.68	5.44
4	3.05	3.99
5	4.10	5.21
6	7.05	10.26
7	6.60	13.91
8	4.79	18.53
9	7.39	7.91
10	7.30	4.85
11	11.78	11.10
12	3.90	3.74
13	26.00	94.03
14	67.48	94.03
15	17.04	41.70

11- One sample: test on a single proportion

$H_0: p = p_0$

$H_1: p \neq p_0$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

For a two tailed test the critical region is $z < -z\alpha/_2 \, or \; z > z\alpha/_2$.

For the one sided alternative $p < p_0$, the critical region is $< -z_\alpha$ and for alternative $p > p_0$, the critical region is $z > z_\alpha$

Example: a commonly prescribed drug for relieving nervous tension is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance.

12- Two samples: test on two proportions

$H_0: p_1 = p_2$

$H_1: p_1 \neq p_2$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(\frac{1}{n_1} + \frac{1}{n_2})}} \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

For a two tailed test the critical region is $z < -z\alpha/_2 \, or \; z > z\alpha/_2$.

For the one sided alternative $p_1 < p_2$, the critical region is $< -z_\alpha$ and for alternative $p_1 > p_2$, the critical region is $z > z_\alpha$

Example: A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. The construction site is within the town limits, and for this reason many voters in the county believe that the proposal will pass because of the large proportion of town voters who favor the construction. To determine if there is a significance difference in the proportions of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county residents favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use a 0.05 level of significance.

13- One sample tests concerning variance

$H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 \neq \sigma_0^2$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

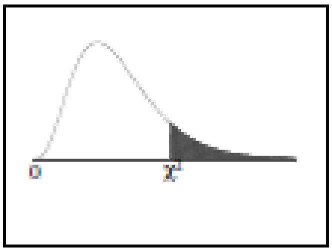
Where n is the sample size, s^2 is the sample variance, and σ_0^2 is the value of σ^2 given by the null hypothesis.

If H_0 is true, χ^2 is a value of chi-square distribution with v=n-1 degrees of freedom.

For a two tailed test, the critical region is $\chi^2 < {\chi^2}_{1-\alpha/2}$ or $\chi^2 > {\chi^2}_{\alpha/2}$. For the one-sided alternative$\sigma^2 < \sigma_0^2$, the critical region is $\chi^2 < {\chi^2}_{1-\alpha}$, and for the one sided alternative $\sigma^2 > \sigma_0^2$, the critical region is $\chi^2 > {\chi^2}_{\alpha}$.

Example: A manufacturer of car batteries claims that the life of the company's batteries is approximately normally distributed with a standard deviation equal to 0.9 year. If a random sample of 10 from these batteries has a standard deviation of 1.2 years, do you think that $\sigma > 0.9$ year? Use a 0.05 level of significance.

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

14- Two sample tests concerning variances

Test the equality of the variances σ_1^2 and σ_2^2 of two populations.

$H_0: {\sigma_1}^2 = \sigma_2^2$

$H_1: {\sigma_1}^2 < \sigma_2^2, {\sigma_1}^2 > \sigma_2^2, {\sigma_1}^2 \neq \sigma_2^2$

For independent random samples of size n_1 and n_2, respectively, from the two populations, the f-value for testing ${\sigma_1}^2 = \sigma_2^2$ is the ratio $f = \frac{s_1^2}{s_2^2}$

F-distribution with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom.

The critical regions of the one-sided alternatives ${\sigma_1}^2 < \sigma_2^2$, ${\sigma_1}^2 > \sigma_2^2$ are, respectively, $f < f_{1-\alpha(v_1,v_2)}$ and $f > f_{\alpha(v_1,v_2)}$. For the two-sided alternative ${\sigma_1}^2 \neq \sigma_2^2$, the critical region is $f < f_{1-(\alpha/2)(v_1,v_2)}$ or $f > f_{(\alpha/2)(v_1,v_2)}$

Example: An experiment was performed to compare the abrasive wear of two different laminated materials. 12 pieces of material 1 were tested by exposing each piece to a machine measuring wear. 10 pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 with a sample standard deviation of 5.

Check whether or not that the two unknown population variances were equal. Use a 0.05 level of significance.

Critical values of the F distribution for $\alpha = 0.05$; $P\{F > F_{.05}(k_1,k_2)\} = 0.05$

Denomin df (k_2)	Numerator Degrees of Freedom (k_1) 1	2	3	4	5	6	7	8	9	10	11	12	15	20	24	30	40	60	100
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.0	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.0
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.9	8.8	8.8	8.8	8.8	8.7	8.7	8.7	8.6	8.6	8.6	8.6	8.6
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.41
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.71
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.76
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.59
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.46
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.35
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.26
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.19
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.12
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.07
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.02
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.98
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.94
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.91
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.88
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.85
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.82
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.80
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.78
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.76
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.74
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.73
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.70
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.59
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.87	1.78	1.74	1.69	1.63	1.58	1.52
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.48
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88	1.79	1.70	1.65	1.60	1.54	1.48	1.43
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.77	1.68	1.63	1.57	1.52	1.45	1.39

15- Goodness of fit test

The test is based on how good a fit we have between the frequency of occurrence of observations in an observed sample and the expected frequencies obtained from the hypothesized distribution.

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Where χ^2 is a value of chi-squared distribution with v=k-1 degrees of freedom.

$o_i: Observed\ frequencies, for\ the\ ith\ cell$

$e_i: Expected\ frequencies, for\ the\ ith\ cell$

$H_0: A\ good\ fit$

$H_1: A\ poor\ fit$

If $\chi^2 > \chi^2_\alpha$, H_0 will be rejected.

Example: to illustrate the tossing of a die, we hypothesize that the die is honest, which is equivalent to testing the hypothesis that the distribution of outcomes is the discrete uniform distribution $p(x) = {}^1/_6$, $x = 1,2, \dots ,6$

Suppose that the die is tossed 120 times and each outcome is recorded. Theoretically, if the die is balanced, we would expect each face to occur 20 times. The results are given as the following.

Face	1	2	3	4	5	6
Observed	20	22	17	18	19	24
Expected	20	20	20	20	20	20

16- Test for independency (Categorical Data)

The chi-squared test procedure discussed in the previous part can also be used to test the hypothesis of independence of two variables of classification.

H_0: $Two\ variable\ are\ independent$

H_1: $Two\ variable\ are\ dependent$

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \qquad expected\ frequency = \frac{column\ total * row\ total}{grand\ total}$$

If $\chi^2 > {\chi^2}_{\alpha}$ with $v = (r-1)*(c-1)$ degrees of freedom. Reject the null hypothesis of independence at the α level of significance; otherwise, fail to reject the null hypothesis.

Example: suppose that we wish to determine whether the opinions of the voting residents of the state of Illinois concerning a new tax reform are independent of their levels of income. Members of a random sample of 1000 registered voters from the state of Illinois are classified as to whether they are in a low, medium, or high income bracket and whether or not they favor the tax reform. The observed frequencies are presented in table below, which is known as a contingency table.

	Income level			
Tax Reform	Low	Medium	High	Total
For	182	213	213	598
Against	154	138	110	402
Total	336	351	313	1000

$$e_{For/Low} = \frac{336 * 598}{1000} = 200.9$$