

Stat 401, section 10.1 Single Factor ANOVA

notes by Tim Pilachowski

In chapter 8 we conducted hypothesis tests in which we compared a single sample's mean or proportion to some hypothesized value. Chapter 9 expanded this to comparing two means or proportions to each other to determine whether they were different ($\Delta_0 = 0$) or whether the difference was some specified value ($\Delta_0 = \text{some number other than } 0$). Now in chapter 10, we turn to comparing more than two populations using an **analysis of variance**, or ANOVA.

A single-factor ANOVA focuses on a comparison of more than two population or treatment means.

The usual scenarios are:

- (a) different populations: sample data collected to compare a shared characteristic
- (b) same population: sample groups given different treatments and compared.

Notation:

I = the number of populations or treatments being compared

μ_1 = the mean of population 1 or the true average response when treatment 1 is applied

\vdots

μ_I = the mean of population I or the true average response when treatment I is applied.

The letters X and Y were used in two-sample problems (chapter 9) to differentiate the observations in one sample from those in the other. Because this is cumbersome for three or more samples, the usual practice is to use a single letter with two subscripts. The first subscript identifies the sample number, corresponding to the population or treatment being sampled, and the second subscript denotes the position of the observation within that sample.

$X_{i,j}$ = the random variable that denotes the j th measurement taken from the i th population, or the measurement taken on the j th experimental unit that receives the i th treatment.

$x_{i,j}$ = the observed value of the j th measurement taken from the i th population/treatment when the experiment is performed.

While the text uses X_{ij} and x_{ij} when both i and j are single-digit numbers, for the Lectures I'll always have a comma.

For the purposes of section 10.1, all sample sizes will be equal, i.e. $n_1 = n_2 = \dots = n_I = J$.

The individual sample means will be denoted by random variables $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_I$, where $\bar{X}_i = \frac{\sum_{j=1}^J X_{i,j}}{J}$.

Note: The text uses a notation $x_i = \sum_{j=1}^J x_{i,j}$, so that calculated $\bar{x}_i = \frac{x_i}{J}$.

The dot in place of the second subscript signifies that we have added over all values of that second subscript ($j = 1, 2, \dots, J$) while holding the first subscript value (i) fixed.

(A parallel is the means of X and Y in the joint probability distributions of chapter 5.)

In a similar fashion, we can calculate the variance of each sample: $(S_i)^2 = \frac{\sum_{j=1}^J (X_{i,j} - \bar{X}_i)^2}{J-1}$.

The average of all I times J observations, called the **grand mean**, is denoted by $\bar{X}_{..} = \frac{\sum_{i=1}^I \sum_{j=1}^J X_{i,j}}{IJ}$.

Note: The text uses a notation $x_{..} = \sum_{i=1}^I \sum_{j=1}^J x_{i,j}$, so that calculated $\bar{x}_{..} = \frac{x_{..}}{IJ}$.

For a single-factor ANOVA, the hypotheses will be

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_I \text{ versus}$$

$$H_a : \text{at least two of the } \mu_i \text{'s are different.}$$

For example, if we had $I = 3$, we would want to fail to reject H_0 only when $\mu_1 = \mu_2 = \mu_3$,

and would want to reject H_0 when $\mu_1 \neq \mu_2 \neq \mu_3$ or $\mu_1 = \mu_2 \neq \mu_3$ or $\mu_1 \neq \mu_2 = \mu_3$ or $\mu_1 = \mu_3 \neq \mu_2$.

A test of these hypotheses requires that we have available a random sample from each population or treatment.

The observed data will be displayed in a rectangular table, in which samples from the different populations appear in different rows of the table, and $x_{i,j}$ is the j th number in the i th row.

Basic assumptions:

The $x_{i,j}$'s within any particular sample are independent (i.e. we have a random sample taken from the i th population or treatment distribution).

Different samples are independent of one another.

Each of the I population or treatment distributions is normal, and each has the same variance σ^2 .

Note that the I sample standard deviations will generally differ somewhat, even when the corresponding population/treatment σ 's are identical. A rough rule of thumb is that if the largest s is not much more than two times the smallest, it is reasonable to assume equal σ^2 's.

If either the normality assumption or the assumption of equal variances is judged implausible, a method of analysis other than an ANOVA must be used.

Preliminary calculations – sums of squares:

$$\text{The total sum of squares (SST)} = \sum_{i=1}^I \sum_{j=1}^J (x_{i,j} - \bar{x}_{..})^2 = \sum_{i=1}^I \sum_{j=1}^J (x_{i,j})^2 - \frac{1}{IJ} (x_{..})^2.$$

$$\text{The treatment sum of squares (SSTr)} = \sum_{i=1}^I \sum_{j=1}^J (\bar{x}_i - \bar{x}_{..})^2 = \frac{1}{J} \sum_{i=1}^I (x_i)^2 - \frac{1}{IJ} (x_{..})^2.$$

$$\text{The error sum of squares (SSE)} = \sum_{i=1}^I \sum_{j=1}^J (x_{i,j} - \bar{x}_i)^2.$$

$$\text{Recall from earlier, } x_i = \sum_{j=1}^J x_{i,j}, \text{ and } x_{..} = \sum_{i=1}^I \sum_{j=1}^J x_{i,j}.$$

The “shortcut” expressions for SST and SSTr are convenient if ANOVA calculations are to be done by hand. In practice, the wide availability of statistical software will usually make this unnecessary.

The fundamental identity relating the three sums of squares is $SST = SSTr + SSE$. The proof is algebraic, and is given in outline form in the text. In practice, we'll calculate SST and SSTr and then use them to find SSE.

Theory:

If H_0 is true, the J observations in each sample come from a normal population distribution with the *same* mean value μ , in which case the sample means $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_I$ should be reasonably close to one another.

The test procedure is based on comparing a measure of differences among the x_i 's (“between-samples” variation) to a measure of variation calculated from *within* each of the samples.

SST is a measure of the total variation in the data – the sum of all squared deviations about the grand mean.

The identity $SST = SSTr + SSE$ says that this total variation can be partitioned into two pieces. SSE measures variation that would be present (within rows) whether H_0 is true or false, and is thus the part of total variation that is *unexplained* by the status of H_0 .

SSTr is the amount of variation (between rows) that *can be explained* by possible differences in the μ_i 's. H_0 is rejected if the explained variation is large relative to unexplained variation.

Method:

Once SSTr and SSE are computed, each is divided by its associated degrees of freedom to obtain a mean square (*mean* in the sense of average).

The **mean square for treatments** (MSTr) = $\frac{SSTr}{I-1}$. The **mean square for error** (MSE) = $\frac{SSE}{I(J-1)}$.

Proposition:

When H_0 is true, $E(\text{MSTr}) = E(\text{MSE}) = \sigma^2$, whereas when H_0 is false, $E(\text{MSTr}) > E(\text{MSE}) = \sigma^2$.

That is, both statistics are unbiased for estimating the common population variance σ^2 when H_0 is true, but MSTr tends to overestimate σ^2 when H_0 is false.

The unbiased nature of MSE, $E(\text{MSE}) = \sigma^2$, is a consequence of the fact that $E(S_i)^2 = \sigma^2$, whether or not H_0 is true. When H_0 is true, each X_i has the same mean value μ and variance σ^2/J , so the “sample variance” of the

\bar{X}_i 's, $\frac{\sum_{i=1}^I (\bar{X}_i - \bar{X}_{..})^2}{I-1}$ estimates $\frac{\sigma^2}{J}$ unbiasedly. Then, multiplying by J gives MSTr as an unbiased estimator of σ^2 .

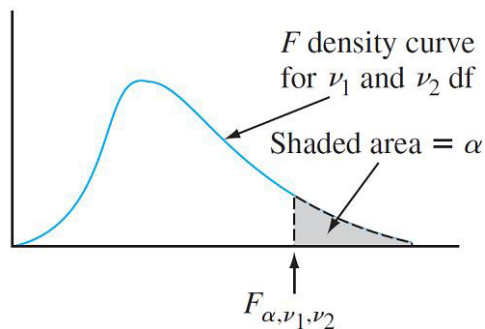
The \bar{X}_i 's tend to spread out more when H_0 is false than when it is true, tending to inflate the value of MSTr when H_0 is false, and we would thus want to reject it.

Test statistic:

Our test statistic, called F , is defined as the ratio of the two mean squares MSTr and MSE: $F = \frac{\text{MSTr}}{\text{MSE}}$.

An F distribution arises in connection with any ratio in which there is one number of degrees of freedom associated with the numerator, and another number of degrees of freedom associated with the denominator. We'll use $\nu_1 = I - 1$ and $\nu_2 = I(J - 1)$ to denote the number of numerator and denominator degrees of freedom, respectively, for our F statistic.

A value of F that greatly exceeds 1, corresponding to an MSTr much larger than MSE, gives us statistical reason to question the assumption that H_0 is true. The appropriate form of the rejection region is therefore $f \geq$ a boundary value c . The value c should be chosen to give $P(F \geq c \text{ when } H_0 \text{ is true}) = \alpha$, the desired significance level. This necessitates knowing the distribution of F when H_0 is true.



In practice, when we have statistical software we will use it to calculate the P -value. When calculating by hand, we'll use Appendix Table A.9 to determine the critical value c for $\alpha = 0.10, 0.05, 0.01, \text{ and } 0.001$. Values of ν_1 are identified with different columns of the table, and the rows are labeled with various values of ν_2 . A calculated test statistic $f \geq c$ implies $p \leq \alpha$, in which case we will reject the null hypothesis.

Our computations will be summarized in a tabular format, called an **ANOVA table**. (See Example A below.)

Example A. A paper in *Measurement and Evaluation in Counseling and Development* (Oct 90, pp. 121–127) discussed a survey instrument called the *Mathematics Anxiety Scale for Children* (MASC). Suppose the MASC was administered to three groups of five sixth graders, with each group having been taught using a different method. Test whether the results of the three methods differ ($\alpha = 0.05$). Data are as follows.

Group 1	67	50	70	60	55	$x_{1.} =$	$\bar{x}_{1.} =$
Group 2	49	32	65	39	43	$x_{2.} =$	$\bar{x}_{2.} =$
Group 3	40	39	41	60	45	$x_{3.} =$	$\bar{x}_{3.} =$
						$x_{..} =$	$\bar{x}_{..} =$

hypotheses:

$$I =$$

$$J =$$

$$v_1 =$$

$$v_2 =$$

$$\sum \sum (x_{i,j})^2 =$$

$$SST =$$

$$SSTr =$$

$$SSE =$$

$$MSTr =$$

$$MSE =$$

$$f =$$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	f
Treatments				
Error				X X
Total			X X	X X

critical value c (from Table A.9) =

conclusion:

Appendix Table A.9 Critical Values for F Distributions

		$\nu_1 = \text{numerator df}$								
ν_2	α	1	2	3	4	5	6	7	8	9
1	0.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.050	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	0.010	4052.20	4999.50	5403.40	5624.60	5763.60	5859.00	5928.40	5981.10	6022.50
	0.001	405,284	500,000	540,379	562,500	576,405	585,937	592,873	598,144	602,284
2	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	0.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.001	998.50	999.00	999.17	999.25	999.30	999.33	999.36	999.37	999.39
3	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.001	167.03	148.50	141.11	137.10	134.58	132.85	131.58	130.62	129.86
4	0.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	0.001	74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.47
5	0.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	0.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	0.001	47.18	37.12	33.20	31.09	29.75	28.83	28.16	27.65	27.24
6	0.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.001	35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69
7	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.001	29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33
8	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.001	25.41	18.49	15.83	14.39	13.48	12.86	12.40	12.05	11.77
9	0.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.001	22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.001	21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96
11	0.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.001	19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.35	8.12
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.001	18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48

Appendix Table A.9 Critical Values for F Distributions

		$\nu_1 = \text{numerator df}$								
ν_2	α	10	12	15	20	25	30	40	50	60
1	0.100	60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.69	62.79
	0.050	241.88	243.91	245.95	248.01	249.26	250.10	251.14	251.77	252.20
	0.010	6055.80	6106.30	6157.30	6208.70	6239.80	6260.60	6286.80	6302.50	6313.00
	0.001	605,621	610,668	615,764	620,908	624,017	626,099	628,712	630,285	631,337
2	0.100	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.47
	0.050	19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48	19.48
	0.010	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.48
	0.001	999.40	999.42	999.43	999.45	999.46	999.47	999.47	999.48	999.48
3	0.100	5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15	5.15
	0.050	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.58	8.57
	0.010	27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.35	26.32
	0.001	129.25	128.32	127.37	126.42	125.84	125.45	124.96	124.66	124.47
4	0.100	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.80	3.79
	0.050	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.70	5.69
	0.010	14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.69	13.65
	0.001	48.05	47.41	46.76	46.10	45.70	45.43	45.09	44.88	44.75
5	0.100	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.15	3.14
	0.050	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.44	4.43
	0.010	10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.24	9.20
	0.001	26.92	26.42	25.91	25.39	25.08	24.87	24.60	24.44	24.33
6	0.100	2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.77	2.76
	0.050	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.75	3.74
	0.010	7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.09	7.06
	0.001	18.41	17.99	17.56	17.12	16.85	16.67	16.44	16.31	16.21
7	0.100	2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.52	2.51
	0.050	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.32	3.30
	0.010	6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.86	5.82
	0.001	14.08	13.71	13.32	12.93	12.69	12.53	12.33	12.20	12.12
8	0.100	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.35	2.34
	0.050	3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.02	3.01
	0.010	5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.07	5.03
	0.001	11.54	11.19	10.84	10.48	10.26	10.11	9.92	9.80	9.73
9	0.100	2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.22	2.21
	0.050	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.80	2.79
	0.010	5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.52	4.48
	0.001	9.89	9.57	9.24	8.90	8.69	8.55	8.37	8.26	8.19
10	0.100	2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.12	2.11
	0.050	2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.64	2.62
	0.010	4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.12	4.08
	0.001	8.75	8.45	8.13	7.80	7.60	7.47	7.30	7.19	7.12
11	0.100	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.04	2.03
	0.050	2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.51	2.49
	0.010	4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.81	3.78
	0.001	7.92	7.63	7.32	7.01	6.81	6.68	6.52	6.42	6.35
12	0.100	2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.97	1.96
	0.050	2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.40	2.38
	0.010	4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.57	3.54
	0.001	7.29	7.00	6.71	6.40	6.22	6.09	5.93	5.83	5.76