STAT 511

Lecture 19: One-way Analysis of Variance (ANOVA)

Devore: Section 10.1-10.3

Prof. Michael Levine

November 19, 2018

Introduction

- The **Analysis of Variance** refers to a collection of experiments and statistical procedures for the analysis of responses from experimental units
- For now, we only study the **single-factor ANOVA** that involves analysis of data obtained from experiments in which more than two treatments have been used. The treatments are differentiated from each other by different **levels** of a single **factor**
- Example: studying the effects of five different brands of gas in automobile engine operating efficiency (mpg) or an experiment to decide whether the color density of fabric specimens depends on the amount of dye used

Single-factor ANOVA: notation

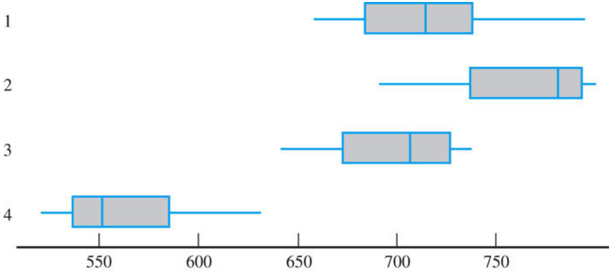
- I - the number of treatments
- $\mu_1, \ldots, \mu_I$ - means of ith population, $i = 1, \ldots, I$
- $H_0 : \mu_1 = \ldots = \mu_I$ vs. H_a : at least two of the μ_is are different
- $X_{i,j}$ - the RV denoting the jth measurement taken from the ith population; $x_{i,j}$ is the observed value of $X_{i,j}$
- It is assumed that $X_{i,j}$ are independent within each sample and the samples are independent of each other

Example

- Several different types of boxes are compared with respect to compression strength (lb)

Table 10.1 The Data and Summary Quantities for Example 10.1

Type of Box	Compression Strength (lb)						Sample Mean	Sample SD
1	655.5	788.3	734.3	721.4	679.1	699.4	713.00	46.55
2	789.2	772.5	786.9	686.1	732.1	774.8	756.93	40.34
3	737.1	639.0	696.3	671.7	717.2	727.1	698.07	37.20
4	535.1	628.7	542.4	559.0	586.9	520.0	562.02	39.87
						Grand mean =	682.50	



(a)

Single-factor ANOVA: notation

- J is the number of observations in each sample I; the data consists of IJ observations
- Sample means are

$$\bar{X}_{i.} = \frac{\sum_{j=1}^{J} X_{ij}}{J}$$

- The **grand means** is

$$\bar{X}_{..} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij}}{IJ}$$

- Sample variances are

$$\frac{\sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2}{J - 1}$$

Single-factor ANOVA: assumptions

- Each $X_{ij} \sim N(\mu_i, \sigma^2)$
- Each of the I populations has the same variance but different means
- If $\frac{\max s_i}{\min s_i} \leq 2$ the equal variance assumption can be assumed to be true
- Normal probability plot of deviations from population means should be used to check the normality assumption

Single-factor ANOVA: test statistic

- Basic idea: compare a measure of differences between $\bar{x}_{i.}$'s to a measure of variation calculated from within each sample
- **Treatment mean square** is

$$MSTr = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$$

- **Error mean square** is

$$MSE = \frac{\sum_{i=1}^{I} S_i^2}{I}$$

- The value of $MSTr$ is affected by the status of H_0 while that of MSE is not

Proposition

- When H_0 is true,

$$E(MSTr) = E(MSE) = \sigma^2$$

- When H_0 is false,

$$E(MSTr) > E(MSE) = \sigma^2$$

F-test

- Consider the F-statistic

$$F = \frac{MSTr}{MSE}$$

- Clearly, when H_0 is not true F has to be large...but what is its distribution?
- Under our assumptions and when H_0 is true, F has an F distribution with $\nu_1 = I - 1$ and $\nu_2 = I(J-1)$ df
- Let f be the observed value of F; the P-value is

$$P(F \geq f | H_0 \text{ is true })$$

which is the area under the $F_{I-1,I(J-1)}$ curve to the right of f

Example

- For strength data, $I = 4$ and $J = 6$ so the df are 3 and $4 * 5 = 20$
- The rejection region is $f \geq F_{0.05,3,20} = 3.10$
- The grand mean is $\bar{x}_{..} = 682.50$, $MSTr = 42,455.86$ and $MSE = 1691.12$
- Thus, $f = \frac{MSTr}{MSE} = 25.09$
- The P-value is less than 0.001

Sums of squares

- The **total sum of squares (SST)** is

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{..})^2$$

- The **treatment sum of squares** is

$$SSTr = \sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{x}_{i.} - \bar{x}_{..})^2$$

- **The error sum of squares** is

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{i.})^2$$

Sums of squares

- Fundamental identity

$$SST = SSTr + SSE$$

- Interpretation: total variation in the data consists of
  1. variation between populations that can be explained by differences in means μ_i
  2. variation that would be present within populations even if H_0 were true
- By definition, $MSTr = \frac{SSTr}{m-1}$, and $MSE = \frac{SSE}{I(J-1)}$.
- Thus, explained variation that is large relative to unexplained corresponds to large values of test statistic F

The **ANOVA table** is

Table 10.2 An ANOVA Table

Source of Variation	df	Sum of Squares	Mean Square	f
Treatments	$I - 1$	SSTr	MSTr = SSTr/$(I - 1)$	MSTr/MSE
Error	$I(J - 1)$	SSE	MSE = SSE/$[I(J - 1)]$	
Total	$IJ - 1$	SST		

Example

- According to the article Evaluating Fracture Behavior of Brittle Polymeric Materials Using an IASCB Specimen (J. of Engr. Manuf., 2013: 133140), researchers have recently proposed an improved test for the investigation of fracture toughness of brittle polymeric materials
- Plexiglas is the material of choice; the test was performed by applying asymmetric three-point bending loads on its specimens
- In one experiment, three loading point locations based on different distances from the center of the specimens base were selected, resulting in the following fracture load data (kN):

						$x_{i\cdot}$
	42 mm:	2.62	2.99	3.39	2.86	11.86
Distance	36 mm:	3.47	3.85	3.77	3.63	14.72
	31.2 mm:	4.78	4.41	4.91	5.06	19.16
						$x_{\cdot\cdot} = 45.74$

Example

- First, test normality and equal variance assumptions - both are satisfied!
- The accompanying ANOVA table is

Source	DF	SS	MS	F	P
Dist	2	6.7653	3.3826	48.58	0.000
Error	9	0.6267	0.0696		
Total	11	7.3920			

Multiple comparisons:Tukey's procedure

- Our task: to control all of the $I(I-1)/2$ intervals possible
- Property:

$$\bar{X}_{i.} - \bar{X}_{j.} - Q_{\alpha,I,I(J-1)}\sqrt{MSE/J} \leq \mu_i - \mu_j \leq \bar{X}_{i.} - \bar{X}_{j.} + Q_{\alpha,I,I(J-1)}\sqrt{MSE/J}$$

  for every $i < j$
- Q_α's are critical values of the Tukey distribution. The result is a collection of confidence intervals with *simultaneous* confidence level $100(1-\alpha)\%$
- We are not interested in lower and upper bounds...but only in whether 0 is included in a given confidence interval or not...Thus
  1. Select α and find $Q_{\alpha,I,I(J-1)}$.
  2. Calculate the margin of error $Q_{\alpha,I,I(J-1)}\sqrt{MSE/J}$
  3. List the sample means in increasing order and underline the pairs that differ by less than the margin of error..

Example

- Five different brands of automobile oil filters are tested. μ_i is the average amount of material captured by brand i filters, $i = 1, \ldots, 5$
- The means are $\bar{x}_{1.} = 14.5$, $\bar{x}_{2.} = 13.8$, $\bar{x}_{3.} = 13.3$, $\bar{x}_{4.} = 14.3$ and $\bar{x}_{4.} = 13.1$

  The ANOVA Table is

Table 10.3 ANOVA Table for Example 10.4

Source of Variation	df	Sum of Squares	Mean Square	f
Treatments (brands)	4	13.32	3.33	37.84
Error	40	3.53	.088	
Total	44	16.85		

Alternative formulation of one-way ANOVA

- The model equation

$$X_{ij} = \mu_i + \varepsilon_{ij}$$

with $E(\varepsilon_{ij}) = 0$ and $V(\varepsilon_{ij}) = \sigma^2$

- An alternative parametrization is

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

where $\alpha_i = \mu_i - \mu$ and $\mu = \frac{1}{I}\sum_{i=1}^{I} \mu_i$

- Now we have $I + 1$ parameters with a constraint $\sum \alpha_i = 0$

Alternative formulation of one-way ANOVA

- The new version of the null hypothesis is

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0$$

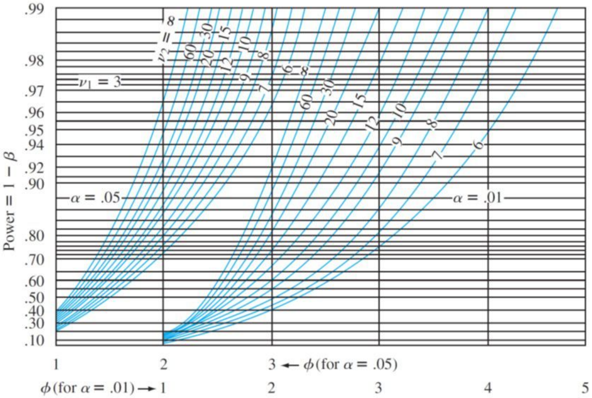
- Under the alternative hypothesis

$$E(MSTr) = \sigma^2 + \frac{J}{I-1} \sum \alpha_i^2$$

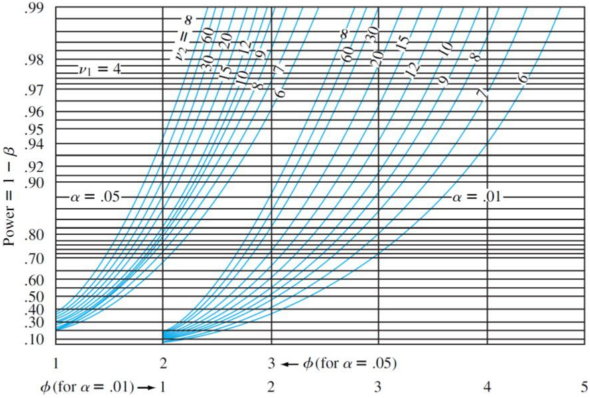
- When H_0 is true, $\sum \alpha_i^2 = 0$ and $E(MSTr) = \sigma^2$
- The larger $\sum \alpha_i^2$ is, the larger the deviation from H_0

Type II Error for F-test

- The distribution of the test statistics under the alternative is a **non-central F distribution**
- Its noncentrality parameter is $\frac{\sum \alpha_i^2}{\sigma^2}$
- The following alternatives provide identical Type II errors: $\alpha_1 = \alpha_2 = -1$,$\alpha_3 = 1$ and $\alpha_4 = 1$ and $\alpha_1 = -\sqrt{2}$, $\alpha_2 = \sqrt{2}$, $\alpha_3 = \alpha_4 = 0$
- The probability of Type II error β is a decreasing function of this parameter



Power = 1 − β
.99
.98
.97
.96
.95
.94
.92
.90
.80
.70
.60
.50
.40
.30
.10
ν₁ = 3
ν₂ = ∞
α = .05
α = .01
1
2
3 ← φ (for α = .05)
φ (for α = .01) → 1
2
3
4
5


.99
.98
.97
.96
.95
.94
.92
.90
.80
.70
.60
.50
.40
.30
.10
Power = 1 − β
$\nu_1 = 4$
$\nu_2 = \infty$
60
30
20
15
12
10
9
8
7
6
α = .05
α = .01
1
2
3 ← ϕ (for α = .05)
ϕ (for α = .01) → 1
2
3
4
5

Example

- The effects of four different heat treatments on yield point (tons/in^2) of steel ingots are to be investigated
- A total of eight ingots will be cast using each treatment. The true standard deviation of yield point for any of the four treatments is $\sigma = 1$
- How likely is it that H_0 will not be rejected at level .05 if three of the treatments have the same expected yield point and the other treatment has an expected yield point that is 1 ton/in^2 greater than the common value of the other three
- In other words, the fourth yield is on average 1 standard deviation above those for the first three treatments

Example

- Thus, $\mu_1 = \mu_2 = \mu_3$, $\mu_4 = \mu_1 + 1$, $\mu = \frac{1}{4}\sum \mu_i = \mu_1 + \frac{1}{4}$
- Therefore, $\alpha_1 = \alpha_2 = \alpha_3 = -\frac{1}{4}$, $\alpha_4 = \frac{3}{4}$
- Compute $\phi^2 = \frac{J}{I}\sum \alpha_i^2/\sigma^2 = \frac{3}{2}$; $\phi = 1.22$
- Df are $\nu_1 = 4 - 1 = 3$ and $\nu_2 = I(J-1) = 28$
- Interpolating visually between $\nu_2 = 20$ and $\nu_2 = 30$ gives power $\approx .47$

Unbalanced design ANOVA and unequal variances

- The total sum of squares is now

$$ST = \sum_{i=1}^{I} \sum_{j=1}^{J_i} (x_{ij} - \bar{x}_{..})^2$$

- The treatment sum of squares is

$$SSTr = \sum_{i=1}^{I} \sum_{j=1}^{J_i} (\bar{x}_{i.} - \bar{x}_{..})^2$$

- The error sum of squares is

$$SSE = \sum_{i=1}^{I} \sum_{j=1}^{J_i} (x_{ij} - \bar{x}_{i.})^2$$

Unbalanced design ANOVA and unequal variances

- The mean sum of squares are $MSTr = \frac{SSTr}{I-1}$ and $MSE = \frac{SSE}{n-I}$
- The rejection region is $f \geq F_{\alpha,I-1,n-I}$

Example

- The article "On the Development of a New Approach for the Determination of Yield Strength in Mg-based Alloys" (Light Metal Age, Oct. 1998: 5153) presented the following data on elastic modulus (GPa)
- The data were obtained by a new ultrasonic method for specimens of a certain alloy produced using three different casting processes
-

									J_i	$x_{i.}$	$\bar{x}_{i.}$
Permanent molding	45.5	45.3	45.4	44.4	44.6	43.9	44.6	44.0	8	357.7	44.71
Die casting	44.2	43.9	44.7	44.2	44.0	43.8	44.6	43.1	8	352.5	44.06
Plaster molding	46.0	45.9	44.8	46.2	45.1	45.5			6	273.5	45.58
									22	983.7	

Example

- Let μ_1, μ_2, μ_3 denote the true average elastic moduli for the three processes
- $H_0 : \mu_1 = \mu_2 = \mu_3$ vs. H_a : at least one of the μ is different
- Df are $I - 1 = 2$ and $n - I = 22 - 3 = 19$
-

Source of Variation	df	Sum of Squares	Mean Square	f
Treatments	2	7.93	3.965	12.56
Error	19	6.00	.3158	
Total	21	13.93		

Multiple comparisons for unequal design ANOVA

- If the imbalance is "mild", the modification of Tukey procedure is used
- Instead of $\frac{1}{J}$, we use the average of the pair $\frac{1}{J_i}$ and $\frac{1}{J_j}$
- In the previous example, $J_1 = J_2 = 8$ and $J_3 = 6$, $I = 3, n - I = 19$, $MSE = .316$
- $Q_{0.05,3,19} = 3.59$and

$$w_{12} = 3.59\sqrt{\frac{3.16}{2}\left(\frac{1}{8} + \frac{1}{8}\right)} = .713$$

- Since $\bar{x}_1 - \bar{x}_2 = .65 < w_{12}$...

Random effects model

- The basic random effects model is

$$X_{ij} = \mu + A_i + \varepsilon_{ij}$$

  where $V(\varepsilon_{ij}) = \sigma^2$ and $V(A_i) = \sigma_A^2$
- A_i and ε_{ij} are normally distributed and independent of one another; $E\,A_i = 0$ is the constraint!
- The null hypothesis are $H_0 : \sigma_A^2 = 0$
- The test statistic is $f = \frac{MSTr}{MSE} \sim F_{I-1,n-I}$ under H_0
- This can be justified by noticing that

$$E(MSTr) = \sigma^2 + \frac{1}{I-1}\left(n - \frac{\sum J_i^2}{n}\right)\sigma_A^2$$