

1-Factor ANOVA Practice Exercises

1) A field study was conducted in the Galapagos Islands where eight individual birds of a single species of finch were collected at random from the finch populations on three randomly selected islands in the Galapagos archipelago. The bill length (in mm) was measured for each bird. Does the bill length of finches vary more among islands or among finches on each island?

2) A tropical trematode parasite that infects humans is characterized by three different life cycle stages within the human bloodstream: stages A, B, and C. Some evidence suggests that the concentration of total proteins (TP) in the blood of humans may differ with the life stage of the parasite, and if so, TP might be used as a simple diagnostic of infection stage. Therefore, the concentration of TP was determined from blood samples obtained from 15 randomly chosen humans infected with the parasite and known to be within each of the three life stages (A – C). Is there a significant difference in blood TP concentration among humans infected by these three different parasite stages? Does TP concentration in humans infected by the parasite in stage A differ from the average of that recorded from those infected by stages B and C (i.e., B+C)?

3) Freshwater wetlands near coastal areas of eastern North America are sometimes inundated with saltwater, particularly at high tide during nor'easter storms, which can affect the survival of some plant species. Therefore, a study was conducted to determine if the survival of a common wetland plant (cattail; *Typha latifolia*) was affected by inundation of seawater for different periods of time. Twenty five cattail plants were placed into individual containers and the containers were then randomly assigned to one of five seawater exposure periods (0 hrs; 4 hrs; 8 hrs; 24 hrs; 48 hrs). The number of weeks that each plant survived was recorded. Is there evidence from these data that survival of cattail plants differed among the five seawater exposure periods? If so, which treatments differ?

DATA SETS

Question 1 Data:

First Island: 6.5, 7.2, 6.9, 15.8, 26.7, 41.7, 21.2, 7.0
 Second Island: 23.0, 15.4, 34.6, 31.8, 58.9, 84.0, 13.3, 47.4
 Third Island: 45.3, 29.1, 35.5, 21.4, 17.4, 31.8, 50.6, 33.2

Question 2 Data

Trt. Group	TP Conc.
A	11
A	10
A	8
A	11
A	7
A	10
A	9
A	7
A	10
A	22
A	21
A	18
A	27
A	31
A	3
B	9
B	18
B	2
B	15
B	20
B	3
B	15
B	23
B	3
B	21
B	15
B	15
B	31
B	125
B	18
C	2
C	2
C	6
C	6
C	2
C	1
C	9
C	0
C	1
C	9
C	6
C	7
C	1
C	6
C	2

Question 3 Data

Trt. Group	Weeks survived
0 hrs	19
0 hrs	32
0 hrs	27
0 hrs	32
0 hrs	33
4 hrs	18
4 hrs	25
4 hrs	28
4 hrs	25
4 hrs	24
8 hrs	17
8 hrs	19
8 hrs	9
8 hrs	12
8 hrs	16
24 hrs	21
24 hrs	21
24 hrs	21
24 hrs	19
24 hrs	19
48 hrs	14
48 hrs	14
48 hrs	12
48 hrs	12
48 hrs	14

Answers for 1-factor ANOVA Practice Exercises

Question 1:

Test required: model II ANOVA & variance components (done by hand)
 Independent variable: Island effect (levels: 3 randomly chosen islands)
 Dependent variable: bill length (mm)
 Number & description of replicates: 8 birds per island
 Statistical hypotheses: Ho: trt groups do not explain significant amount variation in bill length
 Ha: trt groups explain significant amount variation in bill length
 Assumption check: assume all met
 Results: Marginally significant ANOVA; 24.7% variance explained by island effect and 75.3% variance attributable to individual effect (error)
 Conclusion: A significant amount (24.7%) of the variance in bird bill lengths can be attributed to an "island effect", leaving 75.3% of the variance in bill length unexplained.

Tests of Between-Subjects Effects

Dependent Variable: Bill Length (mm)

Source	Type I Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Intercept Hypothesis	20750.520	1	20750.520	19.941	.047	.909
Error	2081.236	2	1040.618 ^a			
ISLAND Hypothesis	2081.236	2	1040.618	3.626	.044	.257
Error	6026.934	21	286.997 ^b			

a. MS(ISLAND)

b. MS(Error)

Variance Estimates

Component	Estimate
Var(ISLAND)	94.203
Var(Error)	286.997

Variance due to island = $(94.203/381.20) \times 100 = 24.7\%$

Error variance = $(286.997/381.20) \times 100 = 75.3\%$

Dependent Variable: LENGTH

Method: ANOVA (Type III Sum of Squares)

Question 2:

Test required:

Independent variable:

Dependent variable:

Number & description of replicates:

Statistical hypotheses:

Assumption check:

Results:

Conclusion:

model I ANOVA (with Helmert contrast)

Parasite life stage (3 levels: stages A, B, C)

total blood protein concentration

blood samples from 15 humans

Ho: no significant differences among means of trt groups

Ha: is a significant difference among means of trt groups

Log transformation is best; all normality & variance assumptions OK if log

ANOVA is highly significant; contrast of trt A vs. mean of B+C is also sig.

Parasite life stage significantly affects the total protein concentration in

human blood (F = 13.917; df = 2, 42; P < 0.0005); however, parasite

life stage A does not result in significantly higher total blood

concentrations (helmert contrast; P = 0.045) when compared to the

mean for stages B and C, although this result is of borderline

significance.

Tests of Normality

TRT	Kolmogorov-Smirnov ^a			Shapiro-Wilk			
	Statistic	df	Sig.	Statistic	df	Sig.	
DATA	a	.295	15	.001	.876	15	.042
	b	.356	15	.000	.536	15	.000
	c	.274	15	.003	.868	15	.031
LOG	a	.210	15	.074	.944	15	.429
	b	.257	15	.009	.892	15	.072
	c	.247	15	.015	.897	15	.087
SROOT	a	.253	15	.011	.928	15	.251
	b	.258	15	.008	.769	15	.001
	c	.232	15	.029	.908	15	.125
POWER	a	.340	15	.000	.763	15	.001
	b	.468	15	.000	.335	15	.000
	c	.301	15	.001	.793	15	.003
INVERSE	a	.195	15	.130	.818	15	.006
	b	.374	15	.000	.708	15	.000
	c	.229	15	.034	.794	15	.003
RANK of DATA	a	.156	15	.200*	.953	15	.567
	b	.278	15	.003	.875	15	.041
	c	.226	15	.038	.903	15	.104

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Test of Homogeneity of Variance

TRT	Based on	Levene	df1	df2	Sig.
		Statistic			
DATA	Based on Mean	2.673	2	42	.081
	Based on Median	1.557	2	42	.223
	Based on Median and with adjusted df	1.557	2	15.919	.241
	Based on trimmed mean	1.670	2	42	.200
	Based on Mean	.896	2	42	.416
LOG	Based on Median	.699	2	42	.503
	Based on Median and with adjusted df	.699	2	33.639	.504
	Based on trimmed mean	.960	2	42	.391
	Based on Mean	1.027	2	42	.367
	Based on Median	.985	2	42	.382
SROOT	Based on Median and with adjusted df	.985	2	22.224	.389
	Based on trimmed mean	.938	2	42	.400
	Based on Mean	4.060	2	42	.024
	Based on Median	1.155	2	42	.325
	Based on Median and with adjusted df	1.155	2	14.127	.343
POWER	Based on trimmed mean	1.555	2	42	.223
	Based on Mean	7.570	2	42	.002
	Based on Median	7.070	2	42	.002
	Based on Median and with adjusted df	7.070	2	23.168	.004
	Based on trimmed mean	7.692	2	42	.001
INVERSE	Based on Mean	1.080	2	42	.349
	Based on Median	.592	2	42	.558
	Based on Median and with adjusted df	.592	2	36.578	.559
	Based on trimmed mean	.896	2	42	.416
	Based on Mean	1.080	2	42	.349
RANK of DATA	Based on Median	.592	2	42	.558
	Based on Median and with adjusted df	.592	2	36.578	.559
	Based on trimmed mean	.896	2	42	.416
	Based on Mean	1.080	2	42	.349
	Based on Median	.592	2	42	.558

Tests of Between-Subjects Effects

Dependent Variable: LOG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.921 ^a	2	1.460	13.917	.000
Intercept	41.820	1	41.820	398.534	.000
TRT	2.921	2	1.460	13.917	.000
Error	4.407	42	.105		
Total	49.148	45			
Corrected Total	7.328	44			

a. R Squared = .399 (Adjusted R Squared = .370)

Contrast Results (K Matrix)

TRT Helmert Contrast		Dependent Variable	
		LOG	
Level 1 vs. Later	Contrast Estimate	.212	
	Hypothesized Value	0	
	Difference (Estimate - Hypothesized)	.212	
	Std. Error	.102	
	Sig.	.045	
	95% Confidence Interval for Difference	5.266E-03	
	Lower Bound	.419	
	Upper Bound		
	Level 2 vs. Level 3	Contrast Estimate	.574
		Hypothesized Value	0
Difference (Estimate - Hypothesized)		.574	
Std. Error		.118	
Sig.		.000	
95% Confidence Interval for Difference		.335	
Lower Bound		.813	
Upper Bound			

Question 3:

Test required: model I ANOVA (with all pairwise multiple comparison)
 Independent variable: Saltwater exposure (5 levels: 0 hrs, 4 hrs, 8 hrs, 24 hrs, 48 hrs)
 Dependent variable: number of weeks that plant survived after saltwater exposure
 Number & description of replicates: 15 individual cattail plants
 Statistical hypotheses: Ho: no significant differences among means of trt groups
 Ha: is a significant difference among means of trt groups
 Assumption check: Square root transformation good: 3 of 5 trt groups normality OK; = variances
 Rank transformation also acceptable
 Results: ANOVA is highly significant regardless of transformation. Multiple comparison results are slightly different depending on transformation; square root provides a more sensitive test.

Conclusion: Exposure to saltwater significantly affects the survival of tail plants ($F = 15.539$; $df = 4, 20$; $P < 0.0005$) in a complex way. Four hr exposure had no effect on survival over that of the control, whereas the other three treatments significantly lowered survival, with the largest differences after 8 and 48 hrs. (THIS IS BASED ON SQRT RESULTS)

Trts: 0hrs 4 hrs 24hrs 8hrs 48hrs

Tests of Between-Subjects Effects

Dependent Variable: SQROOT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	10.294 ^a	4	2.573	15.539	.000
Intercept	489.394	1	489.394	2955.091	.000
TRT	10.294	4	2.573	15.539	.000
Error	3.312	20	.166		
Total	503.000	25			
Corrected Total	13.606	24			

a. R Squared = .757 (Adjusted R Squared = .708)

SQROOT

TRT	N	Subset				
		1	2	3	4	
Tukey HSD ^{a,b}	48 hrs	5	3.6306			
	8 hrs	5	3.7892	3.7892		
	24 hrs	5		4.4931	4.4931	
	4 hrs	5			4.8866	4.8866
	0 hrs	5				5.3227
Sig.			.971	.084	.557	.459
Ryan-Einot-G _a briel-Welsch F	48 hrs	5	3.6306			
	8 hrs	5	3.7892			
	24 hrs	5		4.4931		
	4 hrs	5		4.8866	4.8866	
	0 hrs	5			5.3227	5.3227
Sig.			.860	.318	.244	

RANK of DATA

Tests of Between-Subjects Effects

Dependent Variable: RANK of DATA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	990.700 ^a	4	247.675	16.662	.000
Intercept	4225.000	1	4225.000	284.225	.000
TRT	990.700	4	247.675	16.662	.000
Error	297.300	20	14.865		
Total	5513.000	25			
Corrected Total	1288.000	24			

a. R Squared = .769 (Adjusted R Squared = .723)

TRT	N	Subset		
		1	2	
Tukey HSD ^{a,b}	48 hrs	5	4.80000	
	8 hrs	5	6.70000	
	24 hrs	5	14.60000	
	4 hrs	5	17.80000	
	0 hrs	5	21.10000	
Sig.			.934	.096
Ryan-Einot-G _a briel-Welsch F	48 hrs	5	4.80000	
	8 hrs	5	6.70000	
	24 hrs	5	14.60000	
	4 hrs	5	17.80000	
	0 hrs	5	21.10000	
Sig.			.771	.078

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 14.865.

a. Uses Harmonic Mean Sample Size = 5.000.

b. Alpha = .05.