The null hypothesis for no differences among persons,

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu_8$

is rejected since

$$F(\text{Person}) = MS(\text{Person}) / MS(\text{Error}) = 7.24$$

is greater than $F_{0.95}(7,14) = 2.76$

Module 30: Randomized Block Designs

The first section of this module discusses analyses for randomized block designs. The second part addresses simple repeated measures designs.

Randomized Blocks Designs

The one-way ANOVA is so named because the underlying study design includes, for example k = 4 treatment groups, perhaps with differing numbers of participants in each group. There is no other dimension to the structure.

That is, the only structure is represented within one dimension by the k = 4 treatments. There are circumstances where other dimensions are included.

For example, the k treatments may be included within strata or blocks in an effort to more carefully control for some important sources of variability.

Different age groups, genders, or residents of different communities are examples of such strata or blocks.

When participants within a given block are randomly assigned to one of the treatment groups and this process is repeated for all blocks, the design is called the randomized blocks design.

The resulting two-way structure needs to be taken into account when the data are analyzed.

Blood Pressure Example

The data below represent blood pressure measurements from an experiment involving 4 age groups, each with 3 persons. The 3 persons within each age group were randomly assigned to drugs A, B, and C, with one person per drug.

This was done to keep the drug assignments balanced within age groups.

For this experiment, the major interest is in comparing the three drugs in a manner that provides balance for or controls for possible age effects.

That is, we are interested primarily in hypotheses concerning means for the drugs, but we do have a secondary interest in means for the age groups.

Blood Pressure Data

	Drug			
Age Group	A	B	C	Total
1	70	72	80	222
2	76	84	82	242
3	82	86	84	252
4	90	92	88	270
Total	318	334	334	986

Hypotheses

There are two hypotheses of interest. The one of most importance examines differences among the **Drugs (Treatments),** which can be expressed as:

H_0: $\mu_A = \mu_B = \mu_C$ vs H_1: $\mu_A \neq \mu_B \neq \mu_C$.

The hypothesis that the **Age Group (Block)** means are equal can also be tested. This hypothesis can be written as:

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs H_1: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$

We can examine these hypotheses through the use of an ANOVA procedure appropriate for this **Randomized Blocks** design.

For this design, the blocks—age groups in this case—are selected in order to provide meaningful balance or control for this factor and the random assignment to treatment groups is then done within each block.

ANOVA for Testing Hypotheses

ANOVA

Source	df	SS	MS	F
Treatments	t-1	SS(T)	SS(T)/df(T)	MS(T)/MS(R)
Blocks	b-1	SS(B)	SS(B)/df(B)	MS(B)/MS(R)
Residual	(t-1)(b-1)	SS(R)	SS(R)/df(R)	
Total	n-1	SS(Total)		

For $\alpha = 0.05$, we reject the null hypothesis:

$$H_0: \mu_A = \mu_B = \mu_C ,$$

if

$F(\text{Treatments}) = MS(T)/MS(R) > F_{0.95}[df(T), df(R)]$.

We reject the null hypothesis about block differences,

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 ,$$

if

$F(\text{Blocks}) = MS(B)/MS(R) > F_{0.95}[df(B), df(R)]$.

	Drug							
Age Group	A	B	C	Total	Mean	Sum²/n	Age effect	(Age effect)²n
1	70	72	80	222	74.00	16,428.00	-8.17	200.08
2	76	84	82	242	80.67	19,521.33	-1.50	6.75
3	82	86	84	252	84.00	21,168.00	1.83	10.08
4	90	92	88	270	90.00	24,300.00	7.83	184.08
Total	318	334	334	986	82.17	81,016.33	0.00	401.00

	A	B	C	Total
Mean	79.50	83.50	83.50	82.17
Sum²/n	25,281.00	27,889.00	27,889.00	64,813.67
Drug effect	-2.67	1.33	1.33	0.00
(Drug effect)²n	28.44	7.11	7.11	42.67

Observations squared

	Drug			
Age Group	A	B	C	Total
1	4,900	5,184	6,400	16,484
2	5,776	7,056	6,724	19,556
3	6,724	7,396	7,056	21,176
4	8,100	8,464	7,744	24,308
Total	25,500	28,100	27,924	81,524

The SS calculations are:

$$SS(Total) = 70^2 + 76^2 + \ldots + 84^2 + 88^2 - 986^2/12$$
$$= 81{,}524 - 81{,}016.33$$
$$= 507.67$$

$$SS(T) = 318^2/4 + 334^2/4 + 334^2/4 - 986^2/12$$
$$= 81{,}039 - 81{,}016.33$$
$$= 42.67$$

$$SS(B) = 222^2/3 + 242^2/3 + 252^2/3 + 270^2/3 - 986^2/12$$
$$= 81{,}417.33 - 81{,}016.33$$
$$= 401.00$$

$$SS(R) = SS(Total) - SS(T) - SS(B)$$
$$= 507.67 - 42.67 - 401$$
$$= 64.00$$

ANOVA

Source	df	SS	MS	F
Drug	2	42.67	21.34	1.99
Age Group	3	401.00	133.67	12.47
Residual	6	64.00	10.70	
Total	11	507.67		

F(Drug) = MS(T)/MS(R) = 21.34/10.7 = 1.99,

is not greater than

$$F_{0.95}[df(T), df(R)] = F_{0.95}[2, 6] = 5.14,$$

so H_0: $\mu_A = \mu_B = \mu_C$ should not be rejected.

F(Block) = MS(B)/MS(R) = 133.67/10.7 = 12.47

is greater than

$$F_{0.95}[df(B), df(R)] = F_{0.95}[3, 6] = 4.32$$

so H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ should be rejected.

Looking at models

We can consider a model that characteristizes the individual observations that could be:

$$y_{ij} = \mu + \tau_i + \delta_j + \varepsilon_{ij},$$

where

y_{ij} is the data point for the ith treatment and jth block,

τ_i is the effect for treatment i ,

δ_j is the effect for block j, and

ε_{ij} is the random effect.

Simple Repeated Measures Designs

When a measurement is repeated on each participant so that there are multiple measurements per person, then the resulting dependency of measurements over time on the same person should be considered appropriately when analyses are undertaken.

An example is a study that measured blood pressure levels at several time points for persons assigned to one of k treatment groups.

Recall that we were able to use a paired t-test for testing hypotheses about differences between two time points.

If there are more than two time points, then something else has to be done.

The simplest solution to this situation is to use the procedures outlined for randomized blocks designs discussed above, whereby each participant is considered a block.

The following example includes data for blood pressure measurements over three different time points for each person for a total of eight persons.

Blood pressure measurements at three time points for eight persons

Subject	Baseline	Day 3	Day 7	Total
1	70	73	72	215
2	65	71	69	205
3	68	73	74	215
4	73	75	73	221
5	78	80	76	234
6	67	65	71	203
7	72	72	75	219
8	75	81	74	230
Total	568	590	584	1,742

When the repeated measures are on the same person over time, persons can be treated as "blocks." The randomized block procedure can then be used.

$$SS(Total) = (70)^2 + (65)^2 + \ldots + (74)^2 - (1,742)^2/24$$

$$SS(time) = \frac{568^2}{8} + \frac{590^2}{8} + \frac{584^2}{8} - \frac{1,742^2}{24}$$

$$SS(Persons) = \frac{215^2}{3} + \frac{205^2}{3} + \ldots + \frac{230^2}{3} - \frac{1,742^2}{24}$$

	Blood Pressure						Subject	
Subject	Baseline	Day 3	Day 7	Total	Mean	Sum²/n	Effect	Effect²·n
1	70	73	72	215	71.67	15,408	-0.92	2.52
2	65	71	69	205	68.33	14,008	-4.25	54.19
3	68	73	74	215	71.67	15,408	-0.92	2.52
4	73	75	73	221	73.67	16,280	1.08	3.52
5	78	80	76	234	78.00	18,252	5.42	88.02
6	67	65	71	203	67.67	13,736	-4.92	72.52
7	72	72	75	219	73.00	15,987	0.42	0.52
8	75	81	74	230	76.67	17,633	4.08	50.02
Total	568	590	584	1,742	72.58	126,714	0.00	273.83

	Baseline	Day 3	Day 7	Total
Mean	71.00	73.75	73.00	72.58
Sum²/n	40,328	43,513	42,632	126,472
Time effect	-1.58	1.17	0.42	0.00
(Time effect)²·n	20.06	10.89	1.39	32.33

$1{,}742^2/24 =$ 126,440.17

126,822 - 126,440.17 = 381.83

126,472 - 126,440.17 = 32.33

126,714 - 126,440.17 = 273.83

381.83 - 32.33 - 273.83 = 75.66

	Observations Squared			
Subject	Baseline	Day 3	Day 7	Total
1	4,900	5,329	5,184	15,413
2	4,225	5,041	4,761	14,027
3	4,624	5,329	5,476	15,429
4	5,329	5,625	5,329	16,283
5	6,084	6,400	5,776	18,260
6	4,489	4,225	5,041	13,755
7	5,184	5,184	5,625	15,993
8	5,625	6,561	5,476	17,662
Total	40,460	43,694	42,668	126,822

$$SS(Error) = SS(Total) - SS(Treatments) - SS(Persons)$$

$$SS(Total) = 126{,}822 - 126{,}440.17 = 381.83$$

$$SS(Time) = 126{,}472 - 126{,}440.17 = 32.33$$

$$SS(Persons) = 126{,}714 - 126{,}440.17 = 273.83$$

$$SS(Error) = 381.83 - 32.33 - 273.83 = 75.66$$

ANOVA

Source	df	SS	MS	F
Time	2	32.33	16.17	2.99
Persons	7	273.83	39.11	7.24
Error	14	75.66	5.4	
Total	23	381.83		

The null hypothesis for no time differences,

H_0: $\mu_1 = \mu_2 = \mu_3$,

is accepted since

F(Time) = MS(Time) / MS(Error) = 2.99

is less than $F_{0.95}(2,14) = 3.74$