Homework 1 - BIOS 6643

Dominic Adducci

2023-08-30

Homework 1 Part 1

Question 1 Read the data from the external file and calculate the sample means, standard deviations, and variances at each time point. Make an investigator worthy table of the data.

Table 1: Cortisol Measurements Between Groups

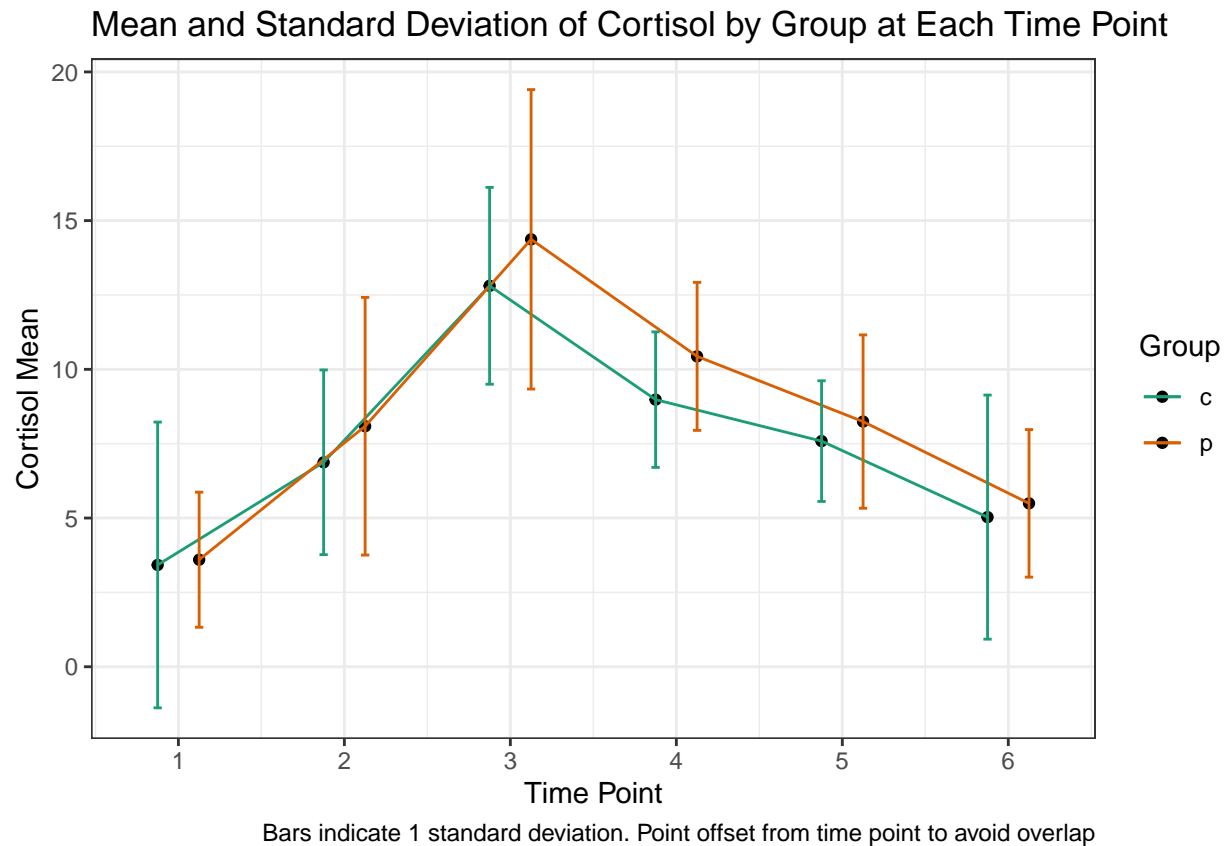|	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6
Mean						
All	3.5107	7.4804	13.5898	9.7086	7.9149	5.2625
Group C	3.4223	6.8739	12.8086	8.9811	7.5864	5.0312
Group P	3.5991	8.0869	14.3710	10.4360	8.2433	5.4938
Standard Deviation						
All	3.7213	3.7820	4.2908	2.4739	2.5079	3.3654
Group C	4.8038	3.1057	3.3092	2.2786	2.0284	4.1036
Group P	2.2710	4.3323	5.0338	2.4885	2.9141	2.4806
Variance						
All	13.8483	14.3036	18.4113	6.1202	6.2898	11.3256
Group C	23.0766	9.6454	10.9506	5.1921	4.1146	16.8394
Group P	5.1576	18.7689	25.3392	6.1925	8.4921	6.1536

Question 2 Interpret the patterns observed in the means and variances for each group in a few sentences.

For group c the mean cortisol measurement increases until time block 3, after which it starts to steadily decrease. This group also has the lowest standard deviation and variance at time block 5. The group trend for standard deviation and variance is that there is a decrease between time block 1 and 2, a slight increase at time block 3, and the a steady decrease between time block 4 and 5, then an increase at time block 6.

For group p the mean cortisol measurement increases until time block 3, after which it starts to steadily decrease. This group also has the lowest standard deviation and variance at time block 1. For this group there is a steady increase in standard deviation and variance from time block 1 to time block 3. There is then a steep decrease at time 4, and then an alternating pattern of increase and decrease in standard deviation and variance between time blocks 4 and 6.

Group c always has a lower mean at each time point compared to group p. Group c has higher standard deviation and variance compared to group p at time block 1 and 6, and lower standard deviation and variance at for all other time blocks.

Question 3 Construct a means and standard deviation graph of the time points with the means connected by a line of a difference color for each group. Interpret the time patterns in the graph. What are your initial thoughts on whether the hypothesis that there are differences in the circadian rhythms between groups will be accepted by your formal analysis?

Mean and Standard Deviation of Cortisol by Group at Each Time Point

Bars indicate 1 standard deviation. Point offset from time point to avoid overlap

Both groups follow the same general trend, where mean cortisol is lowest at time point 1 (22:00-01:50), and steadily increases until time point 3 (06:00-09:50). Both groups steadily decrease in mean cortisol until the final time point (time point 6, 18:00-21:50). This trend follows the circadian rhythm, where cortisol peaks in the middle of the day, and decreases during night.

Between groups the mean cortisol levels are nearly the same for time points 1 and 6. For all other time points in the measurement period group "c" has lower mean cortisol levels compared to group p. The standard deviation for each group varies between time points. For time points 1 and 6 group "c" has a much larger standard deviation compared to group "p". For time points 2, 3, and 5 group "p" has a larger standard deviation. Time point 4 has similar standard deviations between groups.

Mean cortisol levels in general, with the exception of the time point 1 and 6, are lower for group "c" compared to group "p". While this initially seems to suggest that there is a difference in cortisol between women who suffer from depression and those who do not, the standard deviation vary between groups at each time point enough to possible be obscuring what is happening when only looking at means cortisol levels for each time block. Possible explanations for this behavior could be additional variables which may be affecting cortisol levels beyond a diagnosis of depression.

Question 4 Construct a spaghetti plot of each group (one panel for each group). Describe the variation in cortisol levels within a person. Describe the variation in the cortisol levels between people. Which source of variation has more variation? Justify with 1-2 sentences.

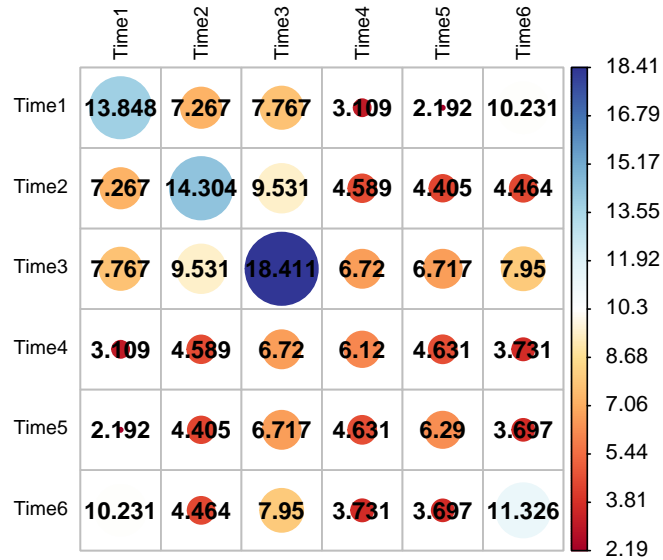## Spaghetti Plots of Mean Cortisol at Each Time Point by Group
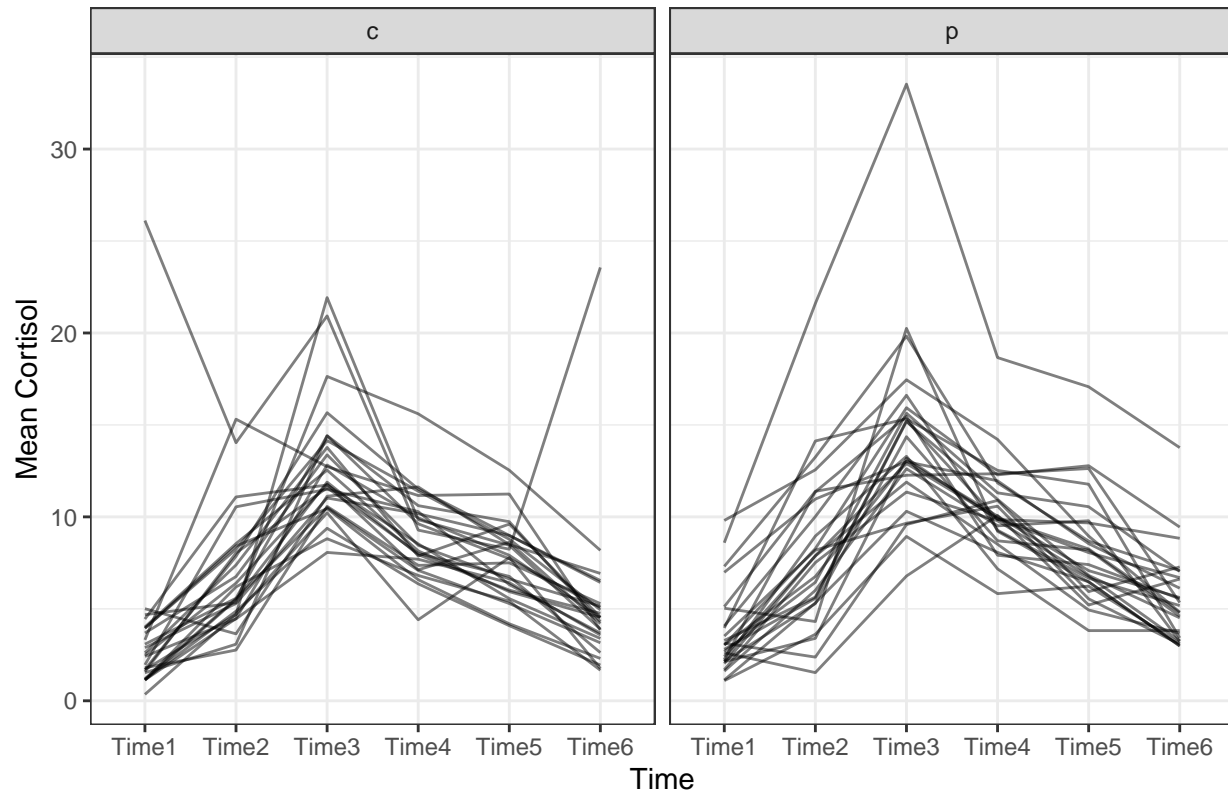
In general the pattern of variation within individuals is the same for both groups, with the middle portion of the day having the highest mean cortisol levels for most individuals, and night/early morning having the lowest. The variation between individuals also changes depending on the time point. For group "c" the general pattern is that the first and last measurements have the least amount of variation between individuals, and for group "p" the variation between individuals is generally lowest at times 1 and 6, with spreading and coalescing of values throughout the measurement period. Additionally, there are a few outliers who may be strongly influencing variance measurements at each time point.
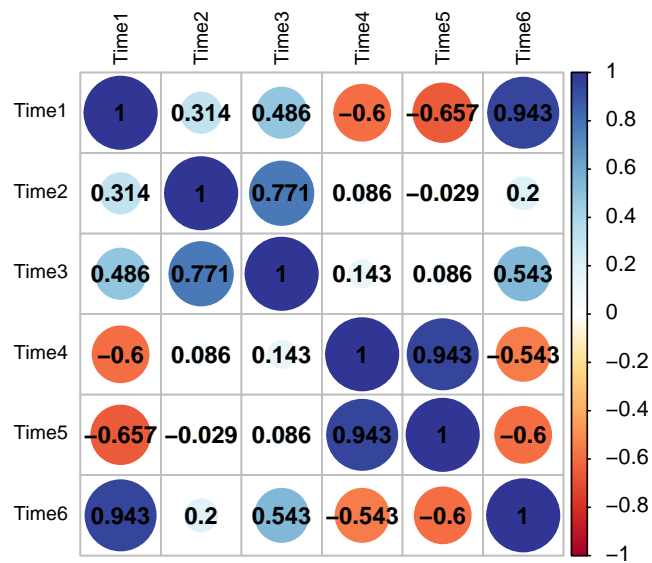
In general there seems to be more variation between subjects compared to within an individual. This is somewhat obscured by the fact that the individuals tend to follow the same trend in mean cortisol levels throughout the day, but overall there are wide variances between individuals during the middle portion of the experimental period.

Question 5   Calculate the 6x6 covariance and correlation matrices for the six time points of cortisol levels.

Covariance Between Cortisol and Time Point

	Time1	Time2	Time3	Time4	Time5	Time6
Time1	13.848	7.267	7.767	3.009	2.192	10.231
Time2	7.267	14.304	9.531	4.589	4.405	4.464
Time3	7.767	9.531	18.411	6.72	6.717	7.95
Time4	3.009	4.589	6.72	6.12	4.631	3.731
Time5	2.192	4.405	6.717	4.631	6.29	3.697
Time6	10.231	4.464	7.95	3.731	3.697	11.326

Correlation Between Cortisol and Time Point

	Time1	Time2	Time3	Time4	Time5	Time6
Time1	1	0.314	0.486	−0.6	−0.657	0.943
Time2	0.314	1	0.771	0.086	−0.029	0.2
Time3	0.486	0.771	1	0.143	0.086	0.543
Time4	−0.6	0.086	0.143	1	0.943	−0.543
Time5	−0.657	−0.029	0.086	0.943	1	−0.6
Time6	0.943	0.2	0.543	−0.543	−0.6	1

Spearman correlation was used due to significant outliers.

Question 6 The covariance matrix is symmetric. Explain why in 1-2 sentences with a mathematical justification.

The covariance matrix is symmetric because the transposition of this matrix is the same as the original matrix. In mathematical terms, using "A" to refer to the covariance matrix:

$$A = A^T$$

Question 7 Verify that the diagonal elements of the covariance matrix are the variances of each time point by comparing your variance-covariance matrix to the table created in problem 1.

Homework 1 Part 2

Test of push