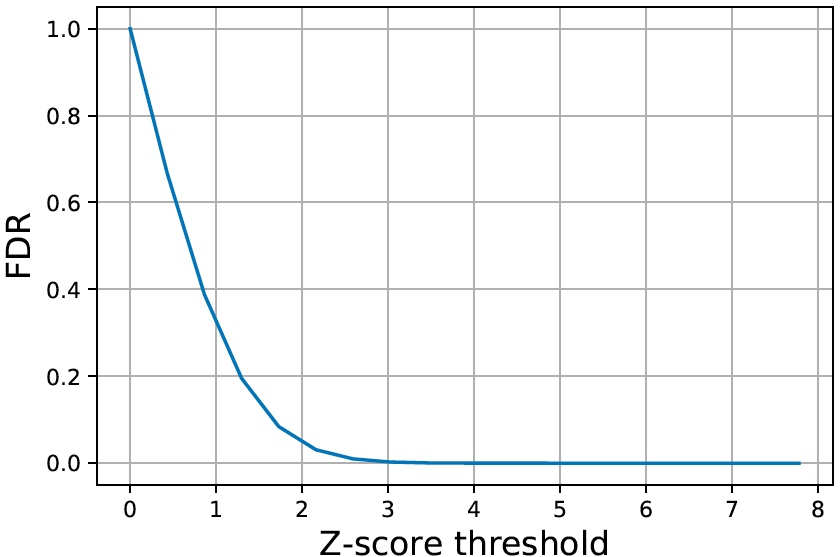
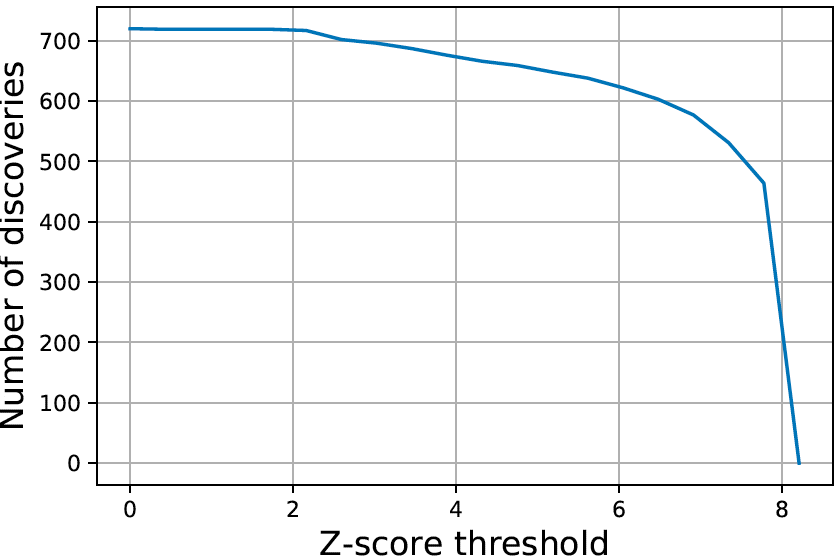
How Is Activity Level Affected by Overall Health?

Naomi Giertych

How much someone moves during the day depends on a variety of factors: job expectations, work schedule, mental health, and physical health. In this analysis, I examine whether an individual’s perception of their overall health influences their activity level using NHANES actigraphy, demographic, and questionnaire data. I regress each measure of activity against the demographic and questionnaire data and perform an F-test for significance of the overall health category. Finally, I compare the false discovery rates (FDR) using the standard measure and local FDR.

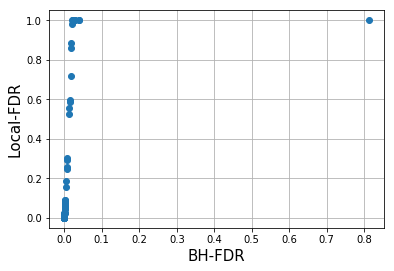
The NHANES actigraphy data contains information on the intensity and duration of movement in 1-minute intervals. I preprocess this data by computing the eigenvalues from a Singular Spectrum Analysis (SSA) for a 12-hour cycle for each of the 7,167 people recruited. The demographic data I use contains information on the respondent’s age, gender, and family income for 10,122 individuals. Finally, I use the overall health response from the questionnaire data for 9,535 individuals; the overall health question asks: “Would you say that in general your health is” and has a scale from 1 to 5 with 1 being excellent and 5 being poor. I remove missing values and non-responses and merge these datasets together; the final dataset contains 5,570 observations.

To determine how an individual’s perception of their overall health influences their activity level, I perform an ANOVA. I regress the activity eigenvalue for each minute of the 12-hour cycle against a person’s age, gender, an interaction between the person’s age and gender, the overall family income, and the self-described overall health measure. Therefore, I have fitted 720 different regression models. I then calculate the F-test statistic between the null model (which excludes the health measure) and the alternative model for each of the models. I calculate the p-value and then convert the p-value to a z-score. Finally, I use the z-scores to estimate the traditional FDR and the local FDR. The traditional FDR is the ratio between the expected number of tests that one would reject given the null hypothesis is true and the number of tests actually rejected. The local FDR is the ratio of the density for the z-scores under the null hypothesis and the the estimated density of the observed z-scores. Below are plots examining the traditional FDR and the number of discoveries against the z-score.

To obtain a small traditional FDR, I would reject all tests with a z-score of 3 or higher, approximately 700 tests. This seems overly optimistic; however, if I were to use a Bonferroni correction, I would reject all tests with a z-score of 3.978 or higher and would reject 699 tests.

Additionally, I compare the traditional FDR to the local FDR. Below is a plot of local FDR against the traditional, Benjamini-Hochberg FDR.



This graph suggests that the traditional FDR is too forgiving compared to the local-FDR. The local-FDR increases quickly even with extremely small values of the traditional FDR. However, even with a traditional FDR of near 0, I can pick a z-score of over 6 and still conclude that the overall health rating is significant in 600 of the regressions.

Overall, I can conclude that how someone perceives their overall health is significantly associated with their physical activity across most eigenvalues of activity level. However, given that the eigenvalues are likely highly correlated further analysis should be performed to form a composite measure of physical activity and compare that with an individual’s perceived health rating.

Sources:

<https://wwwn.cdc.gov/Nchs/Nhanes/2003-2004/PAXRAW_C.htm>

<https://wwwn.cdc.gov/Nchs/Nhanes/2003-2004/DEMO_C.htm>

<https://wwwn.cdc.gov/Nchs/Nhanes/2003-2004/HSQ_C.htm>