**Forecasting (Seasonality)**

Regression with dummy variables

Suppose a seasonality of period *P* exists in the data. We fit the data by multiple linear regression with *P* independent variables. One independent variable is the month. The remaining *P*-1 variables are 0-1 variables representing the seasons (except one). For example, if *P*=4 (W, Sp, Su, F) there are three dummy variables. It does not matter which season is dropped. The dummy variable is equal to 1 if the month is in the specific season, and zero otherwise.

|  |  |  |  |
| --- | --- | --- | --- |
|  | W | Sp | Su |
| W | 1 | 0 | 0 |
| Sp | 0 | 1 | 0 |
| Su | 0 | 0 | 1 |
| F | 0 | 0 | 0 |

Time series analysis

Moving average: A moving average of *P* is the average of *P* months centered at the given month. If *P* is odd, the average includes the month itself and equal number of months before and after. If *P* is even, the average includes the month itself, equal number before and after adding up to *P*-1months and the average of the two extreme values.

Index numbers: The index for a month is the ratio between the sales for this month and the moving average for that month. There are *P* indices, one for each season. We estimate the index for each season by (i) calculating the moving average for each month (except a few first and last months) (ii) calculating the index for each month (iii) calculating the average of all indices for each season (iv) adjusting the averages so that the average of the indices for all *P* seasons is equal to one.

Once the index numbers for each season are established, the adjusted sales are the actual sales divided by the appropriate index. Forecasting based on the adjusted sales is performed (for example, by exponential smoothing). The actual forecast is the adjusted forecast multiplied by the appropriate index.

|  |  |  |
| --- | --- | --- |
|  | Quarter | Sales |
| W | 1 | 72 |
| Sp | 2 | 32 |
| Su | 3 | 88 |
| F | 4 | 24 |
| W | 5 | 80 |
| Sp | 6 | 48 |
| Su | 7 | 72 |
| F | 8 | 32 |