Idősorok

A dekompozíciós modellek lényege, hogy az idősorok négy, egymástól elkülöníthető komponensből tevődnek össze:

- a hosszú távú folyamatokat leíró trendből,
- az ettől szabályos ingadozással eltérő szezonális komponensből,
- a többnyire hosszú távú hullámzást kifejező ciklikus komponensből és
- a véletlen összetevőből.

Megnézem a kapcsolódó epizódot

A lineáris trend egyenlete nagyon egyszerű:

$$\hat{y}_t = \hat{b}_0 + \hat{b}_1 \cdot t$$

A $\hat{b}_0$ és $\hat{b}_1$ paramétereket Excelben vagy bármilyen statisztikai programban néhány kattintással megkapjuk.

Ha kézzel szeretnénk őket kiszámolni, akkor pedig ezekre a normálegyenletekre lesz hozzá szükség:

$$\sum_{t=1}^{n} y_t = n \cdot \hat{b}_0 + \hat{b}_1 \sum_{t=1}^{n} t \qquad \sum_{t=1}^{n} t \cdot y_t = \hat{b}_0 \cdot \sum_{t=1}^{n} t + \hat{b}_1 \sum_{t=1}^{n} t^2$$

Megnézem a kapcsolódó epizódot

A szezonalitást úgy kell elképzelni, hogy az minden nyári szezonban ugyanannyit hozzáad, minden téliben pedig ugyanannyit elvesz a trendvonal által meghatározott értékből.

Pl. ha a négy évszakot vesszük, akkor négy szezonunk van, van egy téli, egy tavaszi, egy nyári és egy őszi, ezért négy szezonalitást kell számolnunk. Más idősorok esetében természetesen ez lehet több is és kevesebb is.

A szezonalitás képlete a következő:

$$s_j = \frac{\sum_{i=1}^{n/p} (y_{ij} - \hat{y}_{ij})}{n/p}$$

A képlet roppant barátságos, de némi magyarázatra szorul. Mindössze arról van szó, hogy minden egyes szezonra átlagoljuk a trendvonal és a tényleges értékek közötti eltéréseket.

Vagyis a képletben p a szezontípusok száma, n pedig az összes szezon száma, y_{ij} jelenti a tényleges értéket, ahol az ij-t úgy kell érteni, hogy az i-edik év j-edik szezonja. $\hat{y}_{ij}$ pedig ennek a trend szerinti megfelelője.

Megnézem a kapcsolódó epizódot

Ha összeadjuk a nyers szezonális eltéréseket, és ezek összege nem nulla, akkor vesszük az átlagukat.

$$\overline{s} = \frac{s_1+s_2+s_3+s_4}{4}$$

És ezt az átlagot mindegyik nyers szezonális eltérésből levonjuk.

$$\tilde{s}_1 = s_1 - \overline{s}$$

$$\tilde{s}_2 = s_2 - \overline{s}$$

$$\tilde{s}_3 = s_3 - \overline{s}$$

$$\tilde{s}_4 = s_4 - \overline{s}$$

Így kapjuk meg a korrigált szezonális eltéréseket

Megnézem a kapcsolódó epizódot

Minden olyan függvényt, ami az y tengelyre szimmetrikus, páros függvénynek hívunk. Ezek a függvények azt tudják, hogy bármely x-re amelyre értelmezve vannak:

$$f(-x) = f(x)$$

Azokat a függvényeket, amelyek az origóra szimmetrikusak, páratlan függvénynek nevezzük. A páratlan függvények úgy működnek, hogy bármely x-re amelyre értelmezve vannak:

$$f(-x) = -f(x)$$

Megnézem a kapcsolódó epizódot

Ha az x különböző pozitív egész kitevős hatványait összeadjuk vagy kivonjuk, akkor polinomokat kapunk.

A polinomfüggvény általános alakja:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots a_1 x + a_0$$

A legmagasabb fokú tag együtthatóját hívjuk főegyütthatónak.

Megnézem a kapcsolódó epizódot

A mozgóátlagolás lényege, hogy az idősor egyes elemeit a körülötte lévő elemek átlagával helyettesítjük, kisimítva ezzel az esetleges erős hullámzásokat.

A mozgóátlagok kiszámolásának képlete páros és páratlan tagú átlagok esetén eltérő.

Ha a tagok száma páratlan:

$$\hat{y}_t = \frac{y_{t-k}+\dots+y_{t-1}+y_t+y_{t+1}+\dots+y_{t+k}}{2k+1}$$

Ha pedig a tagok száma páros:

$$\hat{y}_t = \frac{\frac{y_{t-k}}{2}+\dots+y_{t-1}+y_t+y_{t+1}+\dots+\frac{y_{t+k}}{2}}{2k}$$

Megnézem a kapcsolódó epizódot

Jelentősen csökkenthetjük a normálegyenletek által okozott szenvedéseket, ha az idő múlását jelentő t paramétert úgy adjuk meg, hogy az összege éppen nulla legyen.

Ekkor

$$\sum y_t = n \cdot \hat{b}_0 \qquad \sum t \cdot y_t = \hat{b}_1 \sum t^2$$

[Megnézem a kapcsolódó epizódot]