

4.7. Korreláció és regresszióanalízis

A biológiában, kertészeti kutatásokban stb. fontos lehet több tulajdonságra, jelenségre vagy mennyiségre vonatkozó kísérleti anyagból a vizsgálni kívánt tulajdonságok, jelenségek vagy mennyiségek kapcsolatának meghatározása. A kapcsolat meghatározása annyit jelent, hogy megállapítjuk a mennyiségek egymástól való függésének jellegét és mértékét.

A kapcsolatok egyik legismertebb fajtája a *függvénykapcsolat*, amelyeknél jól definiált összefüggés kapcsolja össze a két vagy több mennyiségre vonatkozó mérési adatokat. (Pl. a fizikából jól ismert, hogy az egyenletesen változó mozgást végző test által megtett út az idő négyzetével arányos: $s = \frac{a}{2} t^2$.)

Gyakran előfordul azonban, hogy a változó mennyiségek között nem teljesen határozott az összefüggés, az egyik változó minden értékéhez a másik változó bizonyos statisztikus sokasága tartozik. A valószínűségi változók közötti kapcsolat általában nem függvényszerű kapcsolat, mert az egyik változó bármely értéke mellett a másik változó értékének létrejöttében általában igen sok más, véletlen hatás is közrejátszik. Ebben az esetben *sztochasztikus* kapcsolatról beszélünk.

A sztochasztikus kapcsolatok vizsgálatának matematikai statisztikai módszere a *korreláció és regresszióanalízis*. (Korreláció: kölcsönösség, kölcsönös viszonylat)

A *korrelációanalízis* a kapcsolat szorosságát vizsgálja.

A *regresszióanalízis* a X valószínűségi változó megfigyelt értékei és egy Y valószínűségi változó várható értéke közti összefüggést megkísérli függvényszerű kapcsolattal leírni. A kapott függvényt *regressziós függvénynek* nevezik.

4.7.1. Korrelációanalízis

A 3.2.4.6. pontban már szoltunk a korrelációs együtthatóról, mint a sztochasztikus kapcsolat erősségének mérésére szolgáló egyik legfontosabb mérőszámról. A 4.2.6. pontban pedig megadtuk annak mintából történő becslését.

Ha a korrelációs együttható közel van az 1 -hez illetve -1 -hez, akkor a két változó közti kapcsolat megközelíti a lineáris függvénykapcsolatot.

Ha X és Y független, a korrelációs együttható értéke $\rho = 0$. Ennek ellenére ρ becslt értéke (r) nullától különböző is lehet.

becslt érték