

Generálj 3 darab egyenként 1000 mintát tartalmazó véletlen oszlopot (változót)! 1 oszlop legyen (0,1) paraméterű normális eloszlású, egy oszlop 1 paraméterű exponenciális eloszlású és egy oszlop eloszlása legyen (0,1)-en egyenletes! Vizsgáld meg egymintás Kolmogorov-Szmirnov próbával 90%-os szignifikancia mellett (10% elsőfajú hibával), hogy a változók tényleg tekinthetőek-e olyan eloszlásúaknak, amilyeneknek generálni akartunk őket! Továbbá kérd le a 3 változó csúcsosságát (kurtosis) és ferdeségét (skewness) és értékelj az eredményt. Megoldásképpen egy a megoldás minden lépését tartalmazó Command Syntacsot és egy az eredményeket tartalmazó és azt szövegesen értékelő pdf fájlt várok.

Segítség:

A változók generálását úgy kell csinálni, ahogy a kockadobást is szimuláltuk, csak a megfelelő RV kezdetű függvényeket kell használni. Függvényekről bővebben olvashattok a belinkelt manual-ban. A Kolmogorov-Szmirnov próbát az Analyze/Nonparametric Statistics/1-Sample K-S menüben éritek el, azt a fület kell bepipálni, amilyen eloszláshoz hasonlítotok. A ferdeség és csúcsosság lekérése az Analyze/ Descriptives több parancsával is lehetséges. Az adatokat ki kell exportálni odt-be, törölni kell a felesleges dolgokat, magyarázatot hozzáírni, majd pdf-be kiexportálni. Figyeljetez arra, hogy minden kiadott parancs benne legyen az általatok küldött command syntacs file-ban.

Tesztek egy kiegészítést a gyakorlaton elhangzottakhoz. Az empirikus ferdeség és empirikus csúcsosság statisztikák mellett találtok egy standard error-nak nevezett számot is. Ez nem más, mint becslése az empirikus ferdeség és empirikus csúcsosság szórásának. Ennek segítségével jobban meg tudjátok ítélni, hogy az elméleti ferdeség és csúcsosság mennyire tekinthető nullának. Ha a kiírt empirikus érték többszöröse a standard error-nak, az arra utal, hogy nem 0 az elméleti ferdeség és csúcsosság.

Megjegyzés:

Majd látni fogjátok, hogy a Kolmogorov-Szmirnov tesztnél van egy kis csalás. Generáltok például standard normális eloszlású valószínűségi változót, de a program nem a standard normálishoz hasonlítja, hanem kiszámolja az empirikus várható értéket és szórást, és az olyan paraméterű normálishoz hasonlítja. Mindezt úgy lehet megfogalmazni, hogy a program egy kicsit módosított Kolmogorov-Szmirnov próbát hajt végre, ami hajlik arra, hogy elfogadjuk a nullhipotézist. Ennek az előnye az, hogy általában lehet tesztelni azt, hogy a változó például normális-e, anélkül, hogy ismernénk a várható értéket és a szórását, hátránya a leírt torzítás. Van a probléma orvoslására korrekciós változat. Még az sem kizárt, hogy a PSPP korrekciósan dolgozik, de nem utal rá semmi. Mindenesetre ezzel a kérdéssel nem kell foglalkoznotok, csak megemlítettem.

Ha bármi kérdésetek van, keressetek nyugodtan! Jó munkát kívánok!