

Házi feladat II.

Papíron megoldandó feladat (természetesen a szükséges kvantilist lehet sőt ajánlott is Excel-ben számolni) (6 pont): Az építőmérnöki vállalkozásunk egy alvállalkozó révén jut 25 kg-os cementes zsákokhoz. A legutóbbi beszállításnál próbaképpen lemértük 8 zsák tömegét. A tömegekre a következők adódtak: 24.5, 25.1, 24.5, 24.9, 25.2, 24.8, 24.3, 24.7. Az alábbi hipotézisről szeretnénk egymintás egyoldali U próbával $\varepsilon = 0,03$ elsőfajú hibavalószínűség mellett dönteni (korábbi tapasztalatok alapján tudjuk, hogy a zsákok tömege normális eloszlást követ 0.2 kg szórással):

H_0 : A cementes zsákok tömegének várható értéke legalább 25 kg

H_1 : A cementes zsákok tömegének várható értéke kisebb, mint 25 kg

Segítség: A t-próba egyoldali módosítását tanultuk. Az u-próba egyoldali módosítását analóg módon kell csinálni.

EXCEL feladat (6 pont): Olvasd be a szimulaltadatok.xlsx-t. Az adathalmazban 3 változó (és egy index változó) 200 megfigyelését látjátok (nomen est omen: az adatok szimuláltak).

1. Számold ki mindhárom változó átlag, korrigált empirikus szórás, ferdeség és csúcsosság statisztikáját. Az órán látott blog alapján számolva, a változók ferdeség illetve csúcsosság standard error-jára (a javasolt egyszerű képlet csak a mintaszámtól függ, ezért adódnak ugyanazok az értékek mindhárom változóra) 0.17 illetve 0.34. Mik ezek az értékek? Ezt felhasználva mely változók elméleti ferdesége és csúcsossága tekinthető 0-nak?
2. Elárulom, hogy a 3 háttérváltozó három különböző eloszlást követ: pontosan egy változó normális, exponenciális illetve egyenletes eloszlású. Az (a) pontban kiszámolt leíró statisztikák alapján mit gondolsz melyik-melyik?

R feladat (8 pont): Olvasd be azezen a linken található adatot az `adat<-read.table("ms212.txt", header=TRUE, sep = '\t')` paranccsal az adatokat. Ezután az `adatszur<-subset(adat,Ran==1)` paranccsal válaszd ki azokat a hallgatókat akiknek testmozgás előtt és után mérték meg a pulzusát.

1. Vizsgáld meg alkalmas hipotézisvizsgálati eszközzel, hogy különbözik-e az edzés előtti illetve utáni pulzus várható értéke?
2. Számold is ki a pulzusváltozás változót!
3. Vizsgáld meg azt alkalmas hipotézisvizsgálati eszközzel, hogy eltér-e az alkoholt rendszeres fogyasztók és alkoholt rendszeresen nem fogyasztók pulzusváltozása!

Megjegyzés: Mivel a normalitásról nincs előzetes információnk illetve a minta nem olyan nagy (nagy minta esetén normalitás nélkül is használható a t-próba), ezért én itt a független kétmintás t-próba nemparaméteres megfelelőjét, a Mann-Whitney tesztet tartom

a legjobb eszköznek. Mivel ezt nem tanítottam, ezért jobb eszköz hiányában használjatok t-próbát! Fontos, hogy a Mann-Whitney teszt ne a várható értékek azonossága a nullhipotézise. Bizonyos interpretációban a nullhipotézis az azonos nagyságrendűség.

4. Teszteld le alkalmas hipotézisvizsgálati eszközzel, hogy a Weight (súly) változó normális eloszlást követ-e!

Kiegészítés: A hipotézisvizsgálatai eszközöket 0.04 előírt elsőfajú hibavalószínűség mellett értékel ki!

Ajánlott beadási határidő: december 7 20 óra (Még este elküldöm a javítást)

Beadási határidő: december 9 20 óra