

Excel zh - 2023

1, (4 pont) A következő linken elérhető kísérlet azt vizsgálta, hogy a hallgatók testsúlya gyarapszik-e az egyetem első évében:

https://dasl.datadescription.com/datafile/freshman-15/?_sfm_methods=Paired+Data&_sfm_cases=4+59943

Az adatok elérhetőek tab-al szeparált txt formátumban. Importáld be az adatokat az Excelbe!

Egyszerű kivonással készítsd el a súlygyarapodás változót majd „nagy mintás U-próbával” vizsgáld meg azt a hipotézist, hogy a különbségváltozó várható értéke 0 (vagyis a matematikai statisztika eszközeivel válaszd meg azt a kérdést, hogy szignifikánsan változott-e a testsúly)! Dolgozz „0.1 elsőfajú hibavalószínűséggel”! A megoldást szöveges is értékel ki!

Magyarázat: Ha a szórás nem ismert, de a háttérváltozó normális akkor U helyett T statisztikát kell számolni és a megfelelő szabadságfokú Student-féle t-eloszlás megfelelő kvantilisét kell használni. A „nagy mintás U-próba” a T és az U próba keveréke: a T statisztikát kell kiszámolni (vagyis a szórás a mintából kell becsülni), de az elfogadási intervallumot úgy kell felvenni ahogy az U próbában csináltuk (vagyis a standard normális eloszlás megfelelő kvantilisét kell használni).

2, (3 pont) Van egy olyan hamis kockánk ami a '6'-os értéket 0.4, míg az 1-5 közti értékek mindegyikét 0.12 valószínűséggel adja. Ötször feldobjuk a kockát és score-t számolunk úgy, hogy a második legkisebb és a legnagyobb adódó értéket összedjük. Az öt dobás 1000 szimulációjára támaszkodva közelítsd annak a valószínűségét, hogy az adódó score legalább 7!

3, (3 pont) Legyenek X_1, X_2, \dots, X_k független standard normális eloszlású valószínűségi változók. Tanultuk, hogy a $Z = \sum_{i=1}^k X_i^2$ négyzetösszeg eloszlását k szabadságfokú khi-négyzet eloszlásnak nevezzük. Az Excel beépített KHINÉGYZET.ELOSZLÁS() függvényét használva készíts egy olyan ábrát amely rögzített szabadságfok mellett kirajzolja az eloszlás sűrűségfüggvényét. Mindezt tedd úgy, hogy a rögzített szabadsági fok változtatható paraméter legyen. Az ábrára második görbeként rajzold rá az $\frac{1}{2}$ paraméterű exponenciális eloszlás sűrűségfüggvényét is (pl az EXP.ELOSZLÁS() függvény használatával). Mely szabadsági fokra egyezik meg a két görbe?

Megjegyzés: A teljes igazság az, hogy a khi-négyzet eloszlás a gamma eloszlás speciális esete.

4, (5 pont)

- a) Legyen X egy $F(x) = \sqrt[4]{x}$ if $0 < x < 1$ (és 0 egyébként) eloszlásfüggvényű valószínűségi változó. Szimulálj 1000 független példányt a valószínűségi változóból! A szimulált mintát használva közelítsd numerikusan X elméleti szórását!

Megjegyzés: Most az elméleti szórás nagyon könnyű kiszámolni. A feladat célja a (c) kérdés előkészítése!

- b) Az 1-es feladatban használt adathalmazban számold ki a végsúly (terminal weight) változó tapasztalati csúcosságát. Értékel ki az eredményt azt figyelembevéve, hogy az órán látott blogbeli (<https://brownmath.com/stat/shape.htm>) képlet 0.291-et ad a sztandard hibára.

- c) Volt arról szó, hogy a fenti blog minden háttérváltozó esetén 0.291-et javasol a ferdeség sztandard hibájára ha a minta 68 elemű. A 68 elemű minta 1000 szimulációjára támaszkodva adj becslést a mintából számolt ferdeség statisztika elméleti szórására akkor amikor a háttérváltozó eloszlásfüggvénye az (a)-beli eloszlásfüggvény (vagyis a feladat szimulációval megvizsgálni, hogy 0.291 mennyire állja meg a helyét akkor amikor a háttérváltozó eloszlásfüggvénye az (a)-beli eloszlásfüggvény)!

Jó munkát kívánok!

Excel zh - 2023

1, (4 pont) A következő linken elérhető kísérlet azt vizsgálta, hogy a hallgatók testsúlya gyarapszik-e az egyetem első évében:

https://dasl.datadescription.com/datafile/freshman-15/?_sfm_methods=Paired+Data&_sfm_cases=4+59943

Az adatok elérhetőek tab-al szeparált txt formátumban. Importáld be az adatokat az Excelbe!

Egyszerű kivonással készítsd el a súlygyarapodás változót majd „nagy mintás U-próbával” vizsgáld meg azt a hipotézist, hogy a különbségváltozó várható értéke 0 (vagyis a matematikai statisztika eszközeivel válaszd meg azt a kérdést, hogy szignifikánsan változott-e a testsúly)! Dolgozz „0.1 elsőfajú hibavalószínűséggel”! A megoldást szöveges is értékel ki!

Magyarázat: Ha a szórás nem ismert, de a háttérváltozó normális akkor U helyett T statisztikát kell számolni és a megfelelő szabadságfokú Student-féle t-eloszlás megfelelő kvantilisét kell használni. A „nagy mintás U-próba” a T és az U próba keveréke: a T statisztikát kell kiszámolni (vagyis a szórás a mintából kell becsülni), de az elfogadási intervallumot úgy kell felvenni ahogy az U próbában csináltuk (vagyis a standard normális eloszlás megfelelő kvantilisét kell használni).

2, (3 pont) Van egy olyan hamis kockánk ami a '6'-os értéket 0.4, míg az 1-5 közti értékek mindegyikét 0.12 valószínűséggel adja. Ötször feldobjuk a kockát és score-t számolunk úgy, hogy a második legkisebb és a legnagyobb adódó értéket összedjük. Az öt dobás 1000 szimulációjára támaszkodva közelítsd annak a valószínűségét, hogy az adódó score legalább 7!

3, (3 pont) Legyenek X_1, X_2, \dots, X_k független standard normális eloszlású valószínűségi változók. Tanultuk, hogy a $Z = \sum_{i=1}^k X_i^2$ négyzetösszeg eloszlását k szabadságfokú khi-négyzet eloszlásnak nevezzük. Az Excel beépített KHINÉGYZET.ELOSZLÁS() függvényét használva készíts egy olyan ábrát amely rögzített szabadságfok mellett kirajzolja az eloszlás sűrűségfüggvényét. Mindezt tedd úgy, hogy a rögzített szabadsági fok változtatható paraméter legyen. Az ábrára második görbeként rajzold rá az $\frac{1}{2}$ paraméterű exponenciális eloszlás sűrűségfüggvényét is (pl az EXP.ELOSZLÁS() függvény használatával). Mely szabadsági fokra egyezik meg a két görbe?

Megjegyzés: A teljes igazság az, hogy a khi-négyzet eloszlás a gamma eloszlás speciális esete.

4, (5 pont)

- a) Legyen X egy $F(x) = \sqrt[4]{x}$ if $0 < x < 1$ (és 0 egyébként) eloszlásfüggvényű valószínűségi változó. Szimulálj 1000 független példányt a valószínűségi változóból! A szimulált mintát használva közelítsd numerikusan X elméleti szórását!

Megjegyzés: Most az elméleti szórás nagyon könnyű kiszámolni. A feladat célja a (c) kérdés előkészítése!

- b) Az 1-es feladatban használt adathalmazban számold ki a végsúly (terminal weight) változó tapasztalati csúcsosságát. Értékel ki az eredményt azt figyelembevéve, hogy az órán látott blogbeli (<https://brownmath.com/stat/shape.htm>) képlet 0.291-et ad a sztandard hibára.

- c) Volt arról szó, hogy a fenti blog minden háttérváltozó esetén 0.291-et javasol a ferdeség sztandard hibájára ha a minta 68 elemű. A 68 elemű minta 1000 szimulációjára támaszkodva adj becslést a mintából számolt ferdeség statisztika elméleti szórására akkor amikor a háttérváltozó eloszlásfüggvénye az (a)-beli eloszlásfüggvény (vagyis a feladat szimulációval megvizsgálni, hogy 0.291 mennyire állja meg a helyét akkor amikor a háttérváltozó eloszlásfüggvénye az (a)-beli eloszlásfüggvény)!

Jó munkát kívánok!