Kidolgozott feladatok a nemparaméteres statisztika témaköréből

A tájékozódást mindenféle színkódok segítik. A feladatok eredeti szövege zöld, a megoldások fekete, a figyelmeztető, magyarázó elemek piros színűek.

1. Nyuszika répatorta-gyára nagyszerűen üzemel, viszont a piacon megjelent egy manufaktúra, amely kézműves répatortákat árusít, amelyek állításuk szerint sokkal finomabbak Nyuszika tortáinál. Ezért Nyuszika közvélemény-kutatást végez. Az utcán sétáló állatoktól megkérdezi, hogy ettek-e már az ő tortáiból, illetve, hogy ettek-e a kézműves tortákból, és kéri, hogy értékeljék ezeket 0-100-as skálán (100-as a legjobb). Nyuszika a kísérlet során nem találkozott olyannal, aki mindkét fajta tortából fogyasztott volna. A kapott pontszámok alább olvashatók:

Nyuszika tortája	78; 66; 75; 85; 73; 59
Kézműves torta	91; 83; 71

Nyuszika azt állítja, hogy nincs különbség a saját tortája és a kézműves torta finomságában.

a) Fogalmazzuk meg, hogy milyen nullhipotézist kíván Nyuszika vizsgálni!

A nullhipotézisnek mindig az eloszlás(ok) valamilyen paraméterére kell vonatkoznia. Itt két eloszlásról beszélhetünk, Nyuszika tortáira adott pontszámok eloszlásáról, illetve a kézműves tortákra adott pontszámok eloszlásáról. Mivel a normális eloszlás korántsem egyértelmű (lásd f) kérdés), így inkább a mediánok egyenlőségét érdemes nézni.

$H_0: \quad \mu_1^e = \mu_2^e$

Ahol μ_1^e Nyuszika tortáira adott pontszámok mediánja, μ_2^e pedig a kézműves tortákra adott pontszámok mediánja.

Természetesen a sokaságok mediánjáról van szó.

A gyakorlatban a javítás során maximális ponttal fogadtam el, ha valaki magyarázat nélkül csak a nullhipotézist írta le. Ezt leginkább azért tettem, mert jellemzően kevés pontok születtek a ZHban. Szerintem egyébként elvárható lenne a magyarázó rész is.

Ha valaki a medián helyett várható értéket írt, 3 pontot kapott a 4-ből.

b) Végezzük el az elemzést kismintás módszerrel!

Rangokat kell rendelni a pontszámokhoz, kapcsolt rangok nincsenek.

A kézműves torták kapják a 9; 7; 3-as rangszámokat, Nyuszika tortái pedig a 8; 6; 5; 4; 2; 1-es rangokat (a legnagyobb pontszámhoz rendeltem a legnagyobb rangot, természetesen fordítva is lehet).

Ha a nullhipotézis igaz, akkor a rangok teljesen véletlenszerűen oszlanak meg a két csoport között.

Ez az előbbi mondat nem a nullhipotézis, tehát nem jó válasz az a) kérdésre!

Ha ez igaz, akkor összesen $\binom{9}{3} = \frac{9!}{6! \cdot 3!} = 84$ különböző, azonos valószínűségű kombináció lehetséges, így egy adott kombináció valószínűsége $\frac{1}{84} = 0{,}0119$.

Ahhoz, hogy a nullhipotézist elutasíthassuk, az kell, hogy a rangszám-összegek a legszélsőségesebb 5%-ba essenek. Itt most a kézműves torták rangszámait nézem, mert így kevesebb számot kell nézni.

$\frac{0{,}05}{0{,}0119} = 4{,}2$, azaz ha a legszélsőségesebb 2-2 kombinációnál utasítom el a nullhipotézist, akkor az α valamivel 5% alatt lesz, ha a legszélsőségesebb 3-3 kombinációnál utasítom el, akkor viszont már felette.

Nézzük meg, hogy melyek a legszélsőségesebb kombinációk!

Legnagyobb rangszám-összegek: 9; 8; 7-nél 24, 9; 8; 6-nál 23, 9; 8; 5-nél és 9; 7; 6-nál 22.

Legkisebb rangszám-összegek: 1; 2; 3-nál 6, 1; 2; 4-nél 7, 1; 2; 5-nél és 1; 3; 4-nél 8.

Így tehát vagy 6, 7, 23 és 24-es rangszám-összeg esetén utasítom el a nullhipotézist, ekkor $\alpha = \frac{4}{84} = 0{,}048$, vagy 6, 7, 8, 22, 23 és 24-esnél, ekkor $\alpha = \frac{8}{84} = 0{,}095$.

A példában viszont a rangszám-összeg 19, tehát a nullhipotézist mindenképp el kell fogadni.

Nyilván nem vártam el, hogy a ZHban a megoldást ilyen részletességgel fejtsük ki.

Azaz az adatok nem mondanak ellent annak, hogy a két tortafajta egyformán finom.

Itt nem vettem a szakmai választ nagyon szigorúan, megintcsak azért, mert kevés pont jött össze a többi feladatból.

Ezzel a módszerrel nem kapunk p-értéket, de nincs is rá szükség. Ahhoz, hogy p-értéket kapjunk, az kell, hogy leszámoljuk az összes olyan kombinációt, ahol a rangszám-összeg 19, vagy ennél szélsőségesebb. Ez igen hosszadalmas feladat, így ezt senkinek sem javasolnám.

c) Végezzük el az elemzést nagymintás módszerrel!

A nagymintás módszer esetén a rangszám-összeget normális eloszlással közelítjük.

A rangszám-összeg várható értéke, és varianciája szerepelt a képletgyűjteményben.

Itt is a kézműves tortákra írom fel a megoldást, de természetesen lehetne Nyuszika tortáira is.

Várható érték: $E(W) = \frac{n_1(N+1)}{2} = \frac{3 \cdot 10}{2} = 15$

Variancia: $Var(W) = \frac{n_1(N+1)n_2}{12} = \frac{3 \cdot 10 \cdot 6}{12} = 15$

$z_0 = \frac{19-15}{\sqrt{15}} = 1{,}033$

Az $\alpha = 0{,}05$-höz tartozó z-kritikus érték $\pm 1{,}96$, így a nullhipotézist elfogadjuk.

A d) kérdéshez szükség lesz p-értékre is, így azt is leolvashatjuk.

1,033-hoz a táblázatból 0,848 olvasható le, ez a tőle „balra” eső terület, a szélsőségesebb értékekhez tartozó terület 1 – 0,848 = 0,152. Mivel kétoldali próbáról van szó, így a p-érték ennek a duplája, $p = 0{,}304 > 0{,}05$, azaz a nullhipotézist ez alapján is elfogadjuk.

Ez alapján nem utasíthatjuk el azt, hogy a torták finomsága azonos.

Ugyanaz él, mint amit a b) kérdésnél leírtam a szakmai válasszal kapcsolatban.

d) Ellenőrizzük, hogy megalapozott volt-e a nagymintás módszer használata!

Akkor megalapozott, ha a p-érték (amit a nagymintás módszerből kapunk) $\frac{1}{N+1}$ és $\frac{N}{N+1}$ közé esik. Ezek itt 0,1 és 0,9.

Mivel a $p = 0{,}304$ ezek közé esik, így megalapozott volt a nagymintás módszer használata.

e) Mennyiben lenne informatívabb, ha Nyuszika olyan állatokat kérdezett volna meg, akik mindkét fajta tortából fogyasztottak? Miért?

Ebben az esetben páros próbát lehetne végezni, ami esetén az egyének közötti ingadozás nem befolyásolja a próbastatisztikát, így megbízhatóbb (erősebb) következtetésünk lehet.

A kulcsszó a páros próba. Természetesen ezzel ekvivalens, más megfogalmazást is elfogadtam.

f) Használhatnánk-e itt paraméteres próbát? Miért?

A paraméteres próbához az kell, hogy az adatok normális eloszlásból származzanak.

Ha egy adat pontozással jön létre, akkor arra első sorban úgy tekintünk, hogy az sorrendi skálán értelmezett, ami esetén nem is merül fel a normális eloszlás, azaz nem végezhetünk paraméteres próbát.

Alapvetően ezt a megoldást fogadtam el, ezzel az indoklással.

Azonban ilyen sok kategória esetén (0-100), már igen jól közelíthető normális eloszlással a sorrendi skálán mért valószínűségi változó, így akár végezhetünk is paraméteres próbát. Ha valaki ezt ténylegesen így megindokolta, azaz nem csak annyit említett meg, hogy az adatok normális eloszlásból származnak (ami ugye pont, hogy nem jellemző a pontozással keletkezett adatokra), az is megkapta a maximális pontot.

2. A Kisegér és Nagyegér a híreket olvassa az újságban. Azt írják, hogy az Egérkirály vizsgálni kívánja, hogy az alkohol fogyasztása hatással van-e arra, hogy az alattvalói mennyire könnyen szagolják ki a sajtot. Ehhez kísérletet végeznek. Egy darab sajtot rejtenek el a kísérleti területen, majd egyes egerekkel alkoholt itatnak, egyesekkel pedig nem, és vizsgálják, hogy 10 percen belül megtalálják-e a sajtot vagy sem.
Az eredményeket a kísérlet elvégzése után 2x2-es táblázatba gyűjtik, ez látható alább:

	megtalálta a sajtot	nem találta meg a sajtot
alkoholt ivott	25	22
nem ivott alkoholt	33	13

a) Kisegér szerint itt a sorösszeg rögzített, Nagyegér szerint itt az oszlopösszeg rögzített. Melyikuknek van igaza? Miért?

Kisegérnek van igaza, mert a sorösszeg rögzített. Ez a 2x2-es táblázat clinical trial-típusú esete. A kísérlet során meghatározzák, hogy melyik egér fogyaszt alkoholt, és melyik nem.

b) Hogyan kellett volna elvégezni a kísérletet, hogy a másik egérnek legyen igaza? Mennyire reális ez az elképzelés?

Ahhoz, hogy a másik egérnek legyen igaza, az kell, hogy az oszlopösszeg rögzített legyen, a sorösszeg pedig nem. Ez egy case-control típusú vizsgálatot adna ilyenkor.

Ehhez az kell, hogy az adatokat utólagosan elemezzük. Egy lehetséges példa, hogy az egereknek hagyják, hogy ők döntsék el, hogy isznak-e alkoholt vagy sem, majd megkerestetik velük a sajtot. Az eredményeket dokumentálják. Majd ezen dokumentumok közül kiválasztanak valahányat (fix számút), ahol az egér megtalálta a sajtot, és kiválasztanak megint csak fix számút, ahol nem. Ezután csak ezeket az adatokat elemzik. Ez nyilván nem hatékony, nem reális elképzelés.

Nem jó az a válasz, hogy minden egér maga dönti el, hogy iszik-e alkoholt, majd vizsgálják az összes egérre, hogy megtalálta-e a sajtot vagy sem. Ugyanis ilyenkor csak a teljes összeg rögzített, az oszlopösszegek nem. Mellesleg ez sokkal reálisabb elképzelés.

Más, kreatív választ is elfogadtam, ha teljesült az, hogy az oszlopösszeg rögzített. Ezek még kevésbé reális elképzelések voltak.

c) Milyen nullhipotézis írja le azt, amire az Egérkirály kíváncsi?

$H_0: \quad \pi_1 = \pi_2$

Ahol π_1 annak a valószínűsége, hogy az alkoholt fogyasztó egér megtalálja a sajtot, π_2 annak a valószínűsége, hogy az alkoholt nem fogyasztó egér megtalálja a sajtot.

Szerintem itt is elvárható lenne, hogy a megoldásban benne legyen a jelölések értelmezése, de nem vontam le most pontot emiatt.

d) Hogy döntünk a nullhipotézisről? Mit mondhatunk az Egérkirálynak?

Ezt többféle módon lehet megoldani. Használhatunk z-próbát, és χ^2-próbát is. Használhatunk folytonossági korrekciót. A z-próba esetén Wald, vagy score módszert kell használnunk. Nyilván akármelyik megoldás jó, és elég csak egyet elvégezni.

$$\hat{\pi}_1 = \frac{25}{25+22} = 0{,}53, \quad \hat{\pi}_2 = \frac{33}{33+13} = 0{,}72$$

<u>Wald módszerrel:</u>

$$z_0 = \frac{0{,}53-0{,}72}{\sqrt{\frac{0{,}53(1-0{,}53)}{47} + \frac{0{,}72(1-0{,}72)}{46}}} = -1{,}93$$

Ez beleesik az $\alpha = 0{,}05$-hoz tartozó $\pm 1{,}96$ intervallumba, tehát a nullhipotézist elfogadjuk.

<u>score módszerrel:</u>

Ha a nullhipotézis igaz, akkor a valószínűség a marginálisokból számítható:

$$\hat{\pi} = \frac{25+33}{25+22+33+13} = 0{,}62$$

$$z_0 = \frac{0{,}53-0{,}72}{\sqrt{0{,}62(1-0{,}62)\left(\frac{1}{47}+\frac{1}{46}\right)}} = -1{,}88$$

Ez beleesik az $\alpha = 0{,}05$-hoz tartozó $\pm 1{,}96$ intervallumba, tehát a nullhipotézist elfogadjuk.

<u>χ^2-próbával, observed/expected táblázattal</u>

Ha a nullhipotézis igaz, akkor mindkét sorban $\hat{\pi} = 0{,}62$ a valószínűsége az 1. oszlopba kerülésnek. Így a várt (expected) értékek táblázata így néz ki:

$47 \cdot 0{,}62 = 29{,}14$	$47 \cdot 0{,}38 = 17{,}86$
$46 \cdot 0{,}62 = 28{,}52$	$46 \cdot 0{,}38 = 17{,}48$

Így a próbastatisztika:

$$\chi_0^2 = \frac{(25-29{,}14)^2}{29{,}14} + \frac{(22-17{,}86)}{17{,}86} + \frac{(33-28{,}52)^2}{28{,}52} + \frac{(13-17{,}48)^2}{17{,}48} =$$

$$= \frac{17{,}14}{29{,}14} + \frac{17{,}14}{17{,}86} + \frac{20{,}07}{28{,}52} + \frac{20{,}07}{17{,}48} = 0{,}588 + 0{,}960 + 0{,}704 + 1{,}148 = 3{,}4$$

Ez egy 1 szabadsági fokú χ^2-eloszlású valószínűségi változó. Annak ellenére, hogy az eredeti nullhipotézis kétoldali próbát kíván, itt a négyzetre emelés miatt csak az ennél nagyobb értékek tekintendőek ennél szélsőségesebbnek (ugyanúgy, mint az ANOVA-nál), tehát felső kritikus értéket kell nézni.

A táblázat alapján a kritikus érték 3,841, a próbastatisztika ennél kisebb, azaz a nullhipotézist elfogadjuk.

Nem vontam le pontot akkor, ha valaki itt kétoldali kritikust nézett, azaz gyakorlatilag $\alpha = 0{,}025$-ös szinten dolgozott.

A táblázatból nem olvasható ki a p-érték, statisztikai szoftverrel megaphatjuk: p = 0,065

A nullhipotézist mindegyik esetben elfogadtuk, azaz az adatok nem mondanak ellent annak, hogy az alkohollal és alkohol nélkül azonos valószínűséggel találják meg az egerek a sajtot.
Itt a nullhipotézis nem megfelelő megfogalmazása esetén levontam 1 pontot. Például helytelen megfogalmazás, hogy ez alapján azonos valószínűséggel találják meg a sajtot az alkoholt ivó és nem ivó egerek. (Vagy bármi más, amiből nem látszik egyértelműen, hogy a nullhipotézis elfogadása nem bizonyító erejű.)

3. A Bölcs Bagoly a Wald és score módszerekről mesél a tanulni vágyó fiatal állatoknak. A kicsik viszont nem figyeltek oda, így nem értették meg az összefüggéseket, így kérdéseik támadtak. Azonban a Bölcs Bagoly nappal alszik, így ő nem tud válaszolni a kérdésekre. Válaszoljunk mi a kérdésekre!
   a) Milyen kismintás módszerek esetén van értelme Wald és score módszerről beszélni?

A Wald és score módszerek úgy merülnek fel, hogy amikor a binomiális eloszlást normális eloszlással közelítjük, a nevezőben, a variancia kifejezésében megjelenik az ismeretlen valószínűség, és ezt kell valamivel helyettesíteni.
Ennek következtében SEMMILYEN kismintás módszernél nincs értelme ezen módszerekről beszélni, hiszen ott ez a probléma nem merül fel.

   b) Mit helyettesítünk mivel, ha a Wald módszert alkalmazzuk?

A valószínűség valódi értékét (π) a becsült értékkel ($\hat{\pi}$) helyettesítjük.

   c) Mit helyettesítünk mivel, ha a score módszert alkalmazzuk?

A valószínűség valódi értékét (π) a nullhipotézis szerinti feltételezett értékkel (π_0) helyettesítjük.
Alapvetően elvárható lett volna, hogy leírjuk azt is, hogy miért, meg hol kell ezt a helyettesítést meglépni, de ezek nélkül is megadtam a maximális pontot.
A 2. feladat d) részében részletesen látható, hogy pontosan hogyan kell ezeket a módszereket alkalmazni.

   d) Használhatunk-e folytonossági korrekciót, ha a Wald, illetve a score módszert alkalmazzuk?

A Wald és score módszerek használata nem befolyásolja a folytonossági korrekció használhatóságát, illetve azon nagymintás módszerek esetén, ahol felmerül a Wald és score módszerek alkalmazásának szükségessége, ott lehetséges folytonossági korrekciót alkalmazni. Tehát, igen, használhatjuk mindkét esetben.

4. Az Országos Egyszarvú Tanács kíváncsi arra, hogy van-e összefüggés a között, hogy az egyszarvúk sörénye szivárványszínben tündököl-e, illetve hogy a patájuk aranyszínben csillámlik-e, miközben vágtáznak. Ezért kérdőíveket küldtek ki 200 egyszarvúnak, amelyben rákérdeztek ezekre. Összesen 112 válasz érkezett, ezek eredményei az alábbi 2x2-es táblázatban szerepelnek:

		szivárványszínben tündököl?	
		igen	nem
csillámlik?	igen	34	17
	nem	27	34

   a) Milyen típusú 2x2-es táblázat ez? (sorösszeg / oszlopösszeg / teljes összeg rögzített, páros vagy kétmintás) Miért?

A táblázat páros mintát tartalmaz, és a teljes összeg rögzített, mert csak annyit tudtunk, hogy hány egyszarvú küldte meg a válaszát, és ők összetartozó két kérdésre adták meg a választ.

Nem jó megoldás az, hogy „sem az oszlop- sem a sorösszeg nem rögzített", ugyanis ilyen típusúból kétféle táblázat van, ahol a teljes összeg rögzített (ez), illetve az, ahol a teljes összeg sem rögzített. Utóbbi táblázat-típust másképpen kell kiértékelni.

b) Milyen nullhipotézist fogalmazhat meg az OET?

Az ilyen típusú táblázatoknál kétféle szakmai kérdést lehet feltenni, és ezen múlik a kiértékelés módja. Mindenképpen meg kell jegyezni azt, hogy e kettő közül melyik a helyes, azt csak a szakmai kérdés határozza meg, nem a kísérlet elvégzésének módja (ugyanis az ugyanaz mindkét esetben).

A kétféle kérdés, amit itt megfogalmazhatunk a függetlenség, illetve a szimmetria. A függetlenség azt jelenti, hogy egy kombináció valószínűsége megegyezik a marginális valószínűségek szorzatával. A szimmetria azt jelenti, hogy az 1. kérdésre adott egyik válasz valószínűsége megegyezik a másik kérdésre adott valamelyik válasz valószínűségével. (Ez NEM jelenti azt, hogy bármelyik kérdésre adott két válasz valószínűsége megegyezik, azaz 50%.)

A feladat szövege alapján az OET arra volt kíváncsi, hogy van-e összefüggés a két tulajdonság között, azaz itt függetlenségvizsgálatot kell végezni.

Ennek nullhipotézise:

$$H_0: \quad \pi_{ij} = \pi_{i+} \cdot \pi_{+j}$$

Ahol π_{ij} annak a valószínűsége, hogy egy egyszarvú 1. kérdésre az i. és a 2. kérdésre a j. választ adja. π_{i+} annak a valószínűsége, hogy egy egyszarvú az 1. kérdésre az i. választ adja. π_{+j} annak a valószínűsége, hogy egy egyszarvú a 2. kérdésre a j. választ adja.

A magyarázat szükségességére vonatkozó megjegyzés itt is érvényben van.

c) Milyen statisztikai próbát kell végezni? (χ^2 vagy McNemar?)

A függetlenségvizsgálatot χ^2-próbával kell végezni.

A szimmetria vizsgálatát a kérdésből magától értetődő módon McNemar-próbával kellene végezni.

d) Végezzük el a próbát! Milyen következtetésre juthat az OET ez alapján?

Először a marginálisokból kiszámítjuk a következő valószínűségek becslését:

$$\hat{\pi}_{1+} = \frac{34+27}{112} = 0{,}545, \quad \hat{\pi}_{2+} = 1 - 0{,}545 = 0{,}455$$

$$\hat{\pi}_{+1} = \frac{34+17}{112} = 0{,}455, \quad \hat{\pi}_{+2} = 1 - 0{,}455 = 0{,}545$$

Ezek alapján elkészíthetjük a várt (expected) értékek táblázatát:

$0{,}455 \cdot 0{,}545 \cdot 112 = 27{,}78$	$0{,}455 \cdot 0{,}455 \cdot 112 = 23{,}22$
$0{,}545 \cdot 0{,}545 \cdot 112 = 33{,}22$	$0{,}455 \cdot 0{,}545 \cdot 112 = 27{,}78$

Így a próbastatisztika:

$$\chi_0^2 = \frac{(34-27{,}78)^2}{27{,}78} + \frac{(17-23{,}22)}{23{,}22} + \frac{(27-33{,}22)^2}{33{,}22} + \frac{(34-27{,}78)^2}{27{,}78} =$$

$$= \frac{38{,}69}{27{,}78} + \frac{38{,}69}{23{,}22} + \frac{38{,}69}{33{,}22} + \frac{38{,}69}{27{,}78} = 1{,}393 + 1{,}666 + 1{,}165 + 1{,}393 = 5{,}617$$

Ugyanazon megfontolások miatt, mint a 2. feladatnál, ez is 1 szabadsági fokú χ^2-eloszlású valószínűségi változó, és a kritikus érték 3,841. Ennél a próbastatisztika nagyobb, így a nullhipotézist elutasítjuk.

Ahogy a 2. feladatnál, itt sem vontam le pontot akkor, ha valaki nem vette észre, hogy csak felső kritikus van.
A p-érték csak statisztikai szoftverrel számítható: p = 0,018.
Ezek alapján van összefüggés a két tulajdonság között.
A nullhipotézis elutasítása bizonyító erejű. Tehát itt az helytelen megfogalmazás, hogy „az adatok alapján nem tudjuk bizonyítani a függetlenséget".
Ha valaki McNemar-próbát végzett itt, akkor az erre a feladatrészre járó 8 pontból maximum 4-et szerezhetett meg. Azt nem tekintettem külön hibának, így nem is vontam le érte pontot, hogy a McNemar alapján a függetlenségre következtetett.

Minden statisztikai próbát α = 0,05-ös szinten kell elvégezni.

Képletgyűjtemény:

Wilcoxon-Mann-Whitney: $E(W \mid H_0) = \frac{n_1(N+1)}{2}$; $Var(W \mid H_0) = \frac{n_1(N+1)n_2}{12}$ (ha nincs kapcsolt rang)

Wilcoxon matched pairs: $E(W \mid H_0) = 0$; $Var(W \mid H_0) = \sum R'^2$

Binomiális eloszlás: $p(k) = \binom{n}{k} \cdot \pi^k \cdot (1-\pi)^{n-k}$

$E(k) = n\pi$; $Var(k) = n\pi(1-\pi)$

χ^2-próba 2x2 táblázat esetén (O – tapasztalt; E – várt): $\chi_0^2 = \frac{\sum_i (O_i - E_i)^2}{E_i}$

A χ^2-próbának ez a formája olyan szempontból jó, hogy a legáltalánosabb, és relatíve könnyen megjegyezhető.

McNemar: $z_0 = \frac{b-c}{\sqrt{b+c}}$